UNIVERSITÀ DEGLI STUDI DI MILANO

Ph.D. School in Biochemical Sciences – XXXII cycle

CAPTURE-BASED NEXT GENERATION SEQUENCING IMPROVES THE IMMUNOGLOBULIN/T-CELL RECEPTOR CLONAL MARKERS IDENTIFICATION IN ADULT ACUTE LYMPHOBLASTIC LEUKEMIA PATIENTS LACKING MOLECULAR PROBES

Advisor: Prof. Andrea MOSCA

Co-Advisor: Dr. Orietta SPINELLI

Roberta CAVAGNA

R11602

Academic Year 2018-2019

This Ph.D. internship has been conducted at the "Paolo Belli" hematology laboratory, Hematology and Bone Marrow Transplant Unit, ASST Ospedale Papa Giovanni XXIII, Bergamo.

INDEX

1. ABSTRACT

The evaluation of residual leukemia cells during and after treatment (Minimal Residual Disease, MRD) is the strongest prognostic factor in Acute Lymphoblastic Leukemia (ALL). The gold standard for MRD monitoring is the Allele-Specific Oligonucleotide (ASO) quantitative PCR, which requires at diagnosis the identification of Immunoglobulin/T-Cell Receptor (Ig/TCR) rearrangements as clonality markers. ASO primers are designed based on obtained sequences of V-D-J junction regions, unique for each lymphocyte rearrangement, and employed for MRD assessment. Conventional procedure allows the accurate and sensitive detection of low frequencies leukemic cells (10^{-5}, i.e. 1 leukemic cell among 100000 normal lymphocytes), but this approach is laborious and time-consuming, requiring the development of patient-specific assay conditions. Furthermore, in about 5-10% of cases (defined as MRD-unknown), this method fails to identify suitable markers for MRD evaluation. Recently, alternative amplicon-based Next Generation Sequencing (NGS) techniques have been described for clonality assessment and MRD monitoring in lymphoid malignancies. Although highly sensitive, these approaches are similarly based on the amplification by PCR of selected rearrangements, thus suffering from some restrictions of the conventional method. With the aim to develop a simplified and possibly more efficient method, we designed a capture-based NGS panel targeting the coding regions of V, D and J genes in the Ig/TCR loci, allowing the selection of even uncommon rearrangements and the possibility to combine the Ig/TCR rearrangements detection with the analysis of relevant ALL-related gene alterations. The panel has been validated on 10 diagnostic samples from patients enrolled into the NILG-ALL 09/00 clinical trial, formerly characterized for clonality. By NGS approach, we not only identified correctly all the 51 known Ig/TCR rearrangements, but we also detected 24 additional clonalities, successively confirmed, including rare rearrangements characterized by uncommon V-D-J combinations, oligoclonal rearrangements and low represented clones. Upon validation on well-characterized samples, we applied the NGS approach to retrospectively identify Ig/TCR clonal markers at diagnosis in the cohort of MRD-unknown patients within the NILG-ALL 09/00 clinical trial, in which conventional method failed. Overall, we identified 92 clonal rearrangements, recognizing at least one Ig/TCR rearrangement in 20/23 patients (87%). No Ig/TCR rearrangements were identified for three MRD-unknown patients, all T-lineage. Furthermore, we confirmed the 77% of rearrangements defined by NGS with a level ≥5%, using implemented standard PCR methods or specific oligonucleotides designed on NGS sequences. For the remaining rearrangements, the validation failed because the NGS identified low represented clones or oligoclones, that could not be solved by conventional procedure using family consensus primers. For most cases, designed ASO-qPCR assays based on NGS sequences reached suitable sensitivity for MRD assessment and allowed to correctly follow the leukemia clones, anticipating the hematologic relapse.

The majority of patients proved *TP53* mutated in a previous work belonged to our MRD-unknown cohort since no informative Ig/TCR rearrangements were identified for the design of a suitable

patient-specific molecular probe. For these patients, we not only identified at least one Ig/TCR rearrangement by NGS, but we also highlighted the prevalence of incomplete IGH rearrangements. The presence of incomplete IGH rearrangements was investigated in 25 additional *TP53* wild-type B-ALL patients for which these clonal rearrangements have not been tested at diagnosis, demonstrating the higher prevalence of incomplete IGH rearrangements in *TP53* mutated B-ALL patients, compared to the *TP53* wild type counterpart.

Finally, thanks to the prospective application of capture NGS approach, we identified at least one rearrangement in 10/12 (83%) newly diagnosed ALL patients MRD-unknown, detecting overall 40 clonal rearrangements. For two T-lineage patients, no Ig/TCR clonal markers have been recognized. Based on sequences provided by NGS, we designed ASO-qPCR assays for the MRD assessment in those cases for which at least one Ig/TCR clonal marker was identified, obtaining at least one highly sensitive patient-specific probe in 50% of cases. For two patients for which follow-up samples were available, we used the designed probes to proceed with the MRD assessment, thus ensuring the correct allocation of patients in the most adequate risk-category. The capture NGS panel, targeting both Ig/TCR loci and ALL-related genes, was validated on diagnostic material from 7 adult ALL patients formerly evaluated for Ig/TCR rearrangements and prospectively tested on the 12 newly diagnosed ALL patients MRD-unknown. Overall, we identified 51 gene variants in 18/19 (94.7%) analyzed patients. In all the B-ALL patients we identified single variants involving exclusively one gene (*CREBBP*, *FLT3*, *JAK2*, *PAX5*, and *TP53*), while the only T-ALL patient showing a single mutated gene (*PTEN*) harbored simultaneously two mutations. Indeed, in the remaining T-ALL cases, we identified the presence of gene variants involving at least two genes and we observed a statistically significant correlation between the T-lineage and the increase of the mutations number per patient ($p=0.0282$). Of note, from 2 to 5 gene variants were identified in three patients in which no Ig/TCR rearrangements were recognized, thus the risk-category of the patients has not been molecularly determined. Since the Ig/TCR rearrangements represent surrogate for stable, easy to identify leukemia clonal markers, the mutated gene alterations could represent, in these latter cases, a molecular signature for MRD analysis with alternative methods or for targeted therapy.

In conclusion, the developed capture-based NGS assay allowed the identification of clonal rearrangements in most cases in which standard approach failed, detecting also rare, oligoclonal and low represented rearrangements, not isolated by conventional methods. Therefore, NGS can be of high value in patients with rare rearrangements, suitable for MRD evaluation and MRD-driven treatment. Furthermore, the NGS technology proved useful in identifying minor leukemia clones at diagnosis, so anticipating possible unexpected relapses in patients negative for major clones. In summary, our results suggest that the prospective application of this technology could simplify clonality assessment and improve standard assays development for leukemia monitoring, also providing information about the mutational status of ALL-related genes.

2. INTRODUCTION

2.1. *ACUTE LYMPHOBLASTIC LEUKEMIA*

Acute Lymphoblastic Leukemia (ALL) is a clonal hematological malignancy deriving from genetic lesions occurring in B- or T-lymphoid precursors or hematopoietic stem cells. These aberrations result in abnormal differentiation, proliferation and accumulation of clonal lymphoid progenitors (lymphoblasts) in bone marrow, blood, and extra-medullary sites as lymph nodes, liver, central nervous system, bones, and testicles.[1]

In Europe, the estimated overall annual incidence of ALL and lymphoblastic lymphoma is 1,28 per 100.000 individuals, with significant age-related differences.[2] As a matter of fact, ALL has a bimodal distribution, with the first peak occurring in 2-5 years-old children and the second peak occurring in adults aged ≥50 years. The 80% of ALL cases arise in children, representing about 30% of pediatric tumors, and the remaining 20% of ALL cases arise in the adult population, with an aggressive course and poor prognosis. In fact, the 5-year Overall Survival (OS) is age-related and decreases with the increase of age: in pediatric patients, OS is higher than 80%, while in adult patients OS falls to 20-30%. Among adults, about 75% of patients suffer from B-lineage ALL, while only 25% of cases are T-lineage.[3]

According to World Health Organization (WHO), the diagnosis of ALL requires the presence of at least 20% of lymphoblasts in the bone marrow and is performed based on cytomorphology, flow cytometry, and molecular criteria.[4]

2.1.1. *Cytomorphology and Cytochemistry*

The first parameter evaluated for primary diagnosis of ALL is the morphological feature, historically according to the classification proposed by the FAB (French-American-British) group.[5] Based on this classification, it is possible to identify three different morphological variants of ALL (L1, L2, and L3) reported in **Table 1**, while cases with mixed morphology are described as well.

Subtype	Morphology	Incidence
L1	Small cells with regular nuclear shape and homogeneous chromatin, small or absent nucleoli and scanty cytoplasm, usually without vacuoles	25-30%
L2	Large and heterogeneous blasts with irregular nuclear shape (often clefted), heterogeneous chromatin, one or more usually large nucleoli and variable basophilic cytoplasm, often abundant, which may contain vacuoles	65-70%
L3	Moderate-large size and homogeneous blasts with regular and round-oval nuclear shape, fine granular chromatin, one or more prominent nucleoli and moderate basophilic cytoplasm containing prominent vacuoles	5-10%

Table 1. *Morphological classification of Acute Lymphoblastic Leukemia proposed by the French-American-British (FAB) group.*[5]

Nowadays, the FAB classification has been replaced by the WHO classification,[6] describing two principal subtypes of lymphoid malignancies, such as neoplasms derived from B- and T-lymphoid precursors (B- or T-lymphoblastic leukemia/lymphoma) and neoplasms derived from mature B, T or NK cells. According to the WHO classification, three different categories of lymphoblastic leukemia/lymphoma are described as follows:

- B-lymphoblastic leukemia/lymphoma not otherwise specified;
- B-lymphoblastic leukemia/lymphoma with recurrent cytogenetic alterations;
- T-lymphoblastic leukemia/lymphoma.

B-ALL with recurrent genetic abnormalities is further delineated based on the specific chromosomal rearrangement identified. Morphologically, it is difficult to perform the clear distinction between B- and T-ALL and between B-lineage blasts and normal B-lineage precursors.

From the cytochemical point of view, lymphoid blasts do not express myeloperoxidase (MPO), although very low expression of MPO is described rarely in cases of ALL. On the contrary, more than 95% of ALL cases are positive for the nuclear enzyme Terminal deoxynucleotidyl Transferase (TdT), normally negative in Acute Myeloid Leukemia (AML) and chronic leukemia forms. However, ALL cases negative for TdT (mostly T-lineage) are uncommonly reported.

2.1.2. *Immunophenotype*

Since the limitations of diagnostic morphology and cytochemistry in ALL definition and assessment, flowcytometric analysis became the diagnostic gold standard for the proper lineage attribution and sub-classification. The immunophenotype analysis is based on the recognition of the dominant phenotype in leukemia population by monoclonal antibodies (mAb). Indeed, leukemic blasts are characterized by differential expression of cell surface and intracellular markers, providing information about the lineage identity, differentiation, and maturation stages and eventually aberrations. The lymphoblasts expression of a given cell surface antigen had been set to 20% of total cells to be considered positive, and to 10% concerning intracellular markers. Criteria for immunological classification of ALL have been proposed by the European Group for Immunological Characterization of Leukemia (EGIL).[7,8]

Based on immunophenotyping criteria, approximately 75–80% of cases of adult ALL are of B-lineage and 20–25% are T-lineage, although some ALL cases present mixed B- and T-phenotype features, deriving from the transformation of totipotent hematopoietic stem cells, rather than differentiated lymphoid precursors. Moreover, immunophenotype analysis allows defining four subcategories for B-lineage ALL and four subcategories for T-lineage ALL, as described in **Table 2**.

B-ALL											
	FAB	CD10	CD19	CD22	CD34	CD79a	cIgM	Cμ	HLA-DR	sIg	TdT
Pro-B (B-I)	L1, L2	-	+	+	+	+	-	-	+	-	+
Common B (B-II)	L1, L2	+	+	+	+/-	+	-	-	+	-	+
Pre-B (B-III)	L1	+/-	+	+	+	+	+	+	+	-	+
Mature-B (B-IV)	L3	+/-	+	+	-	+	+/-	-	+	+	-
T-ALL											
	FAB	CD1a	CD2	cCD3	CD4	CD5	CD7	CD8	TcR	TdT	
Pro-T (T-I)	L1, L2	-	+/-	+	-	+/-	+	-	+	+	
Pre-T (T-II)	L1, L2	-	+/-	+	+	+	+	+	+	+	
Cortical T (T-III)	L1, L2	+	+	+	+	+	+	+	+	+	
Mature (T-IV)	L1, L2	-	+	+	+/-	+	+	+/-	+	+/-	

Tdt=Terminal deoxynucleotidyl Transferase; cIgM=Cytoplasmic Immunoglobulin M; sIg=surface Immunoglobulin; HLA-DR=Human Leucocyte Antigen-DR isotype; TcR=T-cell Receptor.

Table 2. *Immunological classification of Acute Lymphoblastic Leukemia proposed by the European Group of Immunological Characterization of Leukemia (EGIL).* [7,8]

2.1.3. *Cytogenetics*

Cytogenetic analysis of leukemia cells revealed in about 90% of patients the presence of clonal karyotypic alterations, including numerical (hyperdiploidy and hypodiploidy) and or structural (translocations) chromosomal aberrations.[9] Hyperdiploidy (≥ 47-50 chromosomes) is referred in about 30% of cases and associated to a better outcome, by contrast, hypodiploidy (≤45 chromosomes) is highlighted in about 5% of ALL cases and associated to poor diagnosis. Only 20% of patients carry normal diploid karyotype, belonging to the intermediate-risk category. Pseudodyploid karyotype, with normal chromosome number and structural anomalies, is observed in about 30% of ALL cases. Chromosomal rearrangements creating chimeric fusion genes commonly involve hematopoietic transcription factors, epigenetic modifiers, cytokine receptors, and tyrosine kinases.[1,3] In general, chromosomal alterations are more common in B-lineage ALL compared to T-ALL, in which aberrations occur recurrently in chromosome 14. Translocations most frequently identified in ALL are listed in **Table 3**. Among these, the one with the greatest impact on prognosis and treatment is the translocation t(9;22), the signature of Philadelphia (Ph) chromosome.

B-lineage ALL	Chromosomal translocation	Involved genes	Frequency in Adults (%)
	t(9;22)(q34;p11)	*BCR-ABL1*	25-30
	t(12;21)(p13;q22)	*ETV6-RUNX1 (TEL-AML1)*	<1
	t(4;11)(q21;q23)	*KMT2A-AFF1 (MLL-AF4)*	5-10
	t(1;19)(q23;p13.3)	*TCF3 (E2A)-PBX1*	<5
	t(8;14)(q24;q32)	*IGH-MYC*	1
	t(17;19)(q22;p13)	*TCF3 (E2A)-HLF*	1
	t(11;19)(q23;p13.3)	*KMT2A-MLLT1*	1
	t(9;11)(p21;q23)	*KMT2A-MLLT3*	
T- lineage ALL	-	*TAL1 del*	10
	t(10;14)(q24;q11)	*TLX1-TRD*	15
	t(7;10)(q34;q24)	*TRB-TLX1*	15
	t(5;14)(q35;q32)	*BCL11B-TLX3*	15
	t(1;14)(p32;q11)	*TRA-TAL1*	1-3

Table 3. *Most frequent chromosomal abnormalities in Acute Lymphoblastic Leukemia.*

Lately, a new entity defined as Philadelphia Chromosome (Ph)-like Acute Lymphoblastic Leukemia had been described.[10] Ph-like ALL occurs in about 10% of children and 30% of young adults. This condition identifies a genetic profile similar to the Philadelphia chromosome but without the fusion protein BCR-

ABL1. In the majority of cases, the genetic lesions are deletions of transcription factors essential for B-cell hematopoiesis. Ph-like ALL is generally associated with poor prognosis.[10]

2.1.4. *Diagnosis*

ALL symptoms, consisting of asthenia, mucosal and or cutaneous bleeding, frequent infections and fever, are caused by the expansion of leukemia clonal population in the bone marrow and, secondarily, by infiltration in other tissue and metabolic disorders. Indeed, leukemia lymphoblasts completely invade the bone marrow, thus implicating a dramatic reduction of normal hematopoiesis.[1,3]

Diagnosis of ALL requires the morphological evaluation of the bone marrow smear, followed by immunophenotype analysis, cytogenetics, and molecular assessment. In particular, molecular diagnostics allows to identify gene alterations associated with chromosomal translocations, as well as small insertions and deletions not identifiable with cytogenetic analysis, by which only larger chromosomal alterations are visible, involving entire chromosomes or large portions of them. Furthermore, molecular biology techniques, based on qualitative and quantitative Polymerase Chain Reaction (PCR) reactions, allow investigating also immunoglobulin and T-cell receptor rearrangements.

2.1.5. *Prognosis*

For untreated patients, ALL is lethal, depending on normal hematopoiesis impairment and derived cytopenia and based on circulating lymphoblasts effects on vital organs. Death is due to infective or hemorrhagic causes. Modern intensive therapeutic regimen allows achieving hematologic Complete Remission (CR) in about 75-90% of cases during the induction phase of chemotherapy, while 5-10% of patients do not achieve CR, precociously dying during treatment.[1] Furthermore, 10% of ALL patients do not respond to induction due to chemo-resistance. Unfortunately, less than 50% of adult patients survive five years or more, because patients relapse in more than half of the cases achieving CR. Treatment failure relies not only on drugs' primary resistance but also on the selection of chemo-resistant leukemic cells subpopulation or the acquisition of new mutations by tumor cells, leading the therapeutic inefficacy. Overall, available therapeutic options allow the cure for 20-40% of ALL patients.[1,3]

Adult ALL patients are classified into three possible relapse risk-categories, defying Very High Risk (VHR), High Risk (HR) and Standard Risk (SR) patients. Clinical and biological factors as age and total

white blood cells (WBCs) count at diagnosis are crucial to define ALL prognosis, having an independent and cumulative effect on patients' survival.[1,11] In particular, the increase of age portends a worsening prognosis, considering both Disease Free Survival (DFS) and OS. Indeed, for ALL patients older than 60 years prognosis is poor and the 3-years OS is less than 20%, probably because of their intolerance to intensive chemotherapy schemes, co-morbidities or the strong association with the t(9;22) Philadelphia positivity. Time for CR achievement is also essential for the prognosis definition since patients reaching CR within the accomplishment of the first chemotherapy cycle benefit from good prognosis.[11]

Furthermore, the immunophenotypic and karyotypic analysis provides information about the patients' risk-category allocation. As a matter of fact, the expression of myeloid-related antigens (CD13 and CD33) and staminality antigen CD34 in lymphoblasts is related to poor prognosis. Karyotype anomalies are associated with poor prognosis: the positivity for chromosomal translocations t(1;19) and t(9p), for trisomy 21, trisomy 8, deletions del(9p) or del(6q) defines a group of intermediate-risk or HR patients. Translocations t(9;22) (BCR-ABL1) (ALL Ph+) and t(4;11) (*KMT2A-AFF1*) define a group of VHR patients, for which DFS is less than 25% or even 10% in ALL-Ph+ patients, although Ph-positive patients have had an important OS improvement with the introduction in the induction multiagent chemotherapy phase of novel agents as Tyrosine Kinase Inhibitors (TKIs).[1]

2.1.6. *Therapy and unmet clinical need*

The treatment of ALL, aiming to eradicate leukemia clones and prevent disease recurrence, is provided in three phases:

- induction of remission;
- post-remission consolidation/intensification of therapy;
- remission maintenance.

During the consolidation phase, preventive treatment of the central nervous system is also performed.[11]

The goal of the modern intensive induction chemotherapy is the elimination of leukemia cells and the remission achievement, meaning lymphoblasts in the bone marrow or peripheral blood are not detectable. The consolidation phase aims to control eventual residual tumor cells after the induction, in order to prevent relapses, and consists of consolidation chemotherapy or hematopoietic stem cell transplantation (HSCT).[1,11] Indeed, the intensity of this stage of the treatment depends upon the risk of relapse factors and risks associated with treatment. Patients belonging to the SR group, generally receive consolidation chemotherapy, since there is no evidence of increased survival compared to SR

patients receiving HSCT. On the contrary, an intensified treatment regimen with HSCT is usually reserved for HR patients in first remission or to relapsed patients after the achievement of a second remission. Stem cell transplantation requires the patient reinfusion of hematopoietic stem cells deriving from healthy related or unrelated donor, fully-matched or partially-matched for HLA major histocompatibility complex. This procedure is preceded by treatment of ALL patients with myeloablative conditioning, consisting of very high doses of chemotherapy or total body irradiation (TBI). In patients non-eligible for myeloablative conditioning, reduced-intensity conditioning and nonmyeloablative conditioning are performed before HSCT. Allogeneic transplantation exerts against leukemia cells the GvL (Graft versus Leukemia) effect since donor stem cells mature into immune cells and destroy any remaining leukemia cells after myeloablative conditioning. Unfortunately, the GvL effect is associated with GvHD (Graft versus Host Disease) complication, in which the immune response attacks some patient's own healthy organs causing symptoms including severe skin rash, diarrhea, and liver damage. GvHD complication is avoided in autologous transplantation, in which patients are reinfused with their own stem cells collected during complete remission. Instead, autologous transplantation is associated with the lack of GvL effect and to impairment of *ex vivo* purging efficacy. Finally, remission maintenance (or continuation) therapy is intended to maintain the CR status of patients and performed for two or three years, although its benefit for adults had not been clearly demonstrated.[3,11,12] In any case, relapsed/refractory ALL still represent a significant unmet for both pediatric and adult hematological neoplasms, therefore the exploration of novel targeted therapies is a promising field of research.

2.2. *MINIMAL RESIDUAL DISEASE*

Adult patients affected by ALL are constantly monitored to assess remission or to promptly catch recurrences, considering that relapse occurs in more than 50% of cases,[13] frequently carrying the same leukemia clone of the disease presentation. Intensive chemotherapy regimen, indeed, is not sufficient to eradicate the complete neoplastic population and persistent lymphoblasts are not detectable by cytomorphological analysis, characterized by a limited sensitivity that allows identifying at least 5% of total cells. Therefore, new approaches have been introduced in order to detect residual leukemia level, indicating with the term Minimal Residual Disease (MRD) the low-level disease, which persists and is characterized by the presence of a few malignant phenotype bearing cells not detectable by morphologic criteria.[14] Both in pediatric and adult ALL, MRD has been demonstrated as the most important independent prognostic factor, crucial for the correct risk-category allocation of patients and useful to evaluate the treatment response, to monitor the disease follow-up and eventually to punctually identify relapse **(Figure 1)**.[13,15–20] Indeed, responsive patients showing a rapid leukemia clearance during treatment have been demonstrated to benefit from an excellent outcome, while patients with leukemia persistence at early time points had a poorer outcome since disease burden during therapy reflects the patients' limited chemosensitivity. Furthermore, MRD is not only crucial for the assessment of initial treatment response, but also the subsequent definition of MRD-based risk groups and consequent risk-stratification. Nowadays, since the association between MRD persistence and relapse-risk, residual leukemia assessment has been used to support risk-oriented treatment strategies. Selected MRD decisional time points are established during the induction phase of chemotherapy, in order to classify patients as MRD-positive and MRD-negative. These latter ones, mostly displaying a satisfying response to conventional chemotherapy, do not necessarily need an intensified regimen or HSCT. By contrast, MRD-positive patients are treated with intensified chemotherapy and they are eligible for HSCT. Moreover, the MRD detection after or even during the treatment is a powerful tool for early recognition of impending relapse, predicting the detection of hematologic overt recurrence and offering the possibility to anticipate salvage therapy.

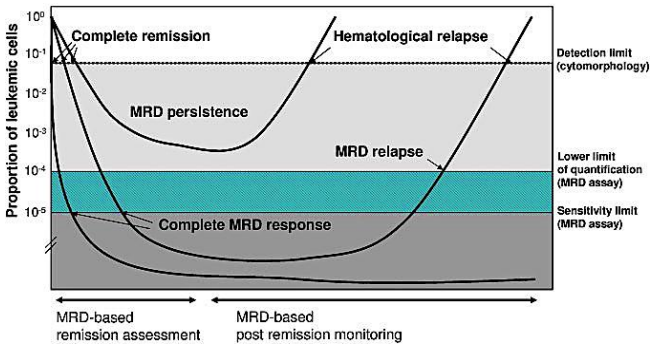

Figure 1. *Proposals for definition of Minimal Residual Disease terms in Acute lymphoblastic Leukemia. Patients showing a rapid leukemia clearance early during treatment benefit from a prolonged MRD negativity, while patients with persistent low-level disease are more prone to molecularly or clinically relapse. Adapted from Brüggemann et al., 2012.*[21]

2.2.1. *Conventional methods for Minimal Residual Disease evaluation*

The MRD techniques need to be sensitive, accurate, reliable, fast, reproducible, broadly applicable, standardized and affordable. Importantly, MRD assays must consider molecular markers stable through-out the disease course. Accurate and sensitive detection of low frequencies of ALL cells, ≤1 leukemia cell among 10000 normal cells (10^{-4} or ≤0.01%),[22,23] requires highly specific markers for discrimination between ALL lymphoblasts and normal leukocytes in bone marrow and peripheral blood and is currently performed by flow cytometry, PCR analysis of breakpoint fusion regions of chromosome aberrations or by PCR analysis of patient-specific junctional regions of rearranged immunoglobulin and T-cell receptor genes. These techniques ensure sufficient sensitivity for MRD detection.

The flow cytometry approach, developed by the EuroFlow Consortium, proposes two 8-color antibody panels for immunophenotyping of hematological malignancies, recognizing Leukemia-associated Aberrant ImmunoPhenotype (LAIP).[24] Although less laborious and faster compared to molecular techniques, the detection of LAIP by flow cytometry allows the identification of one leukemic cell out of 1000 normal lymphocytes (10^{-3}) and requires a quick analysis shortly after sample collection, to avoid cell death. Chromosomal translocations-derived fusion genes, such as *BCR-ABL1*, *ETV6-RUNX1*, *TCF3-PBX1*, *KMT2A-AFF* and *TAL1* deletions (*SIL-TAL1*), and corresponding chimeric transcripts, are suitable markers for MRD detection.[25,26] Indeed, these aberrations are driver events in leukemia pathogenesis, thus leukemia-specific and stables during disease course. Chimeric transcripts are amplified by reverse transcription Rq-PCR and quantified by comparison of serial dilutions of transcripts-containing plasmids or cell lines.[22] Furthermore, in the presence of a fixed genomic breakpoint as in *SIL-TAL1* fusion, a DNA molecular target can be used in T-ALL assessment, avoiding RNA instability issues.

Since only one-third of patients are positive for the presence of chimeric fusion genes,[27] the gold standard for MRD monitoring is the amplification of Immunoglobulin (Ig) and T-Cell Receptor (TCR) genes by Allele-Specific Oligonucleotide quantitative PCR (ASO-qPCR), developed and standardized thanks to EuroMRD consortium efforts.[22,23] This approach requires the identification at diagnosis of Ig/TCR clonal markers by amplification of the V-D-J segments involved in the Ig/TCR gene rearrangements,[28,29] followed by heteroduplex clonality assay and Sanger sequencing. Therefore, the obtained sequences are used to design ASO on V-D-J junction regions, unique for each lymphocyte rearrangement, then employed for MRD evaluation by quantitative PCR in combination with primers and probes recognizing constant regions of the identified rearrangement. The ASO-qPCR approach can reach a reproducible detection limit of up to 10^{-5} (1 leukemic cell among 100000 normal cells)[22,23,30] and ASO-qPCR conditions have to be optimized for each found rearrangement in each patient to find at least two molecular markers with sufficient 10^{-4} sensitivity, following the EuroMRD working group guidelines.

2.2.2. *Lymphocyte ontogeny and Immunoglobulin/T-Cell Receptor gene rearrangements*

During the maturation process, lymphocytes undergo immunophenotypic changes, as well as finely regulated molecular events. In particular, somatic rearrangements occur in Immunoglobulin (Ig) and T-Cell Receptor (TCR) gene loci, in order to increase the antigen receptor diversity in B- and T-lymphocytes, respectively. The deriving vast repertoire of antigen sensor molecules allows the development of a lymphoid clone with an extracellular receptor specific to a given antigen.[31]

Immunoglobulins are present as B-cell membrane receptors as well secreted proteins, while TCRs are heteroduplex molecules exclusively anchored to the external cell surface of T-lineage lymphocytes. Immunoglobulins are composed of four polypeptide subunits: 2 homologous heavy (H) chains and 2 homologous light (L) chains, bonded by disulfide bridges. Each chain is characterized by a constant (C) region, maintaining structural integrity, and an N-terminus variable (V) region, different for each lymphocyte and implicated in the antigen recognition process **(Figure 2A)**. Furthermore, each V region is composed of three CDR (Complementarity Determining Regions) hypervariable regions. CDR3 portion, at the boundaries between V and C segments, is the most variable region, directly involved in the antigen bond. Heavy chain can be of five types (μ, δ, γ, ε and α), defining the class of immunoglobulin IgM, IgD, IgG, IgE, and IgA, respectively, while the light chain can be of two types (λ and κ).

Similarly, TCR proteins are composed of two polypeptide chains, both characterized by a variable (V) portion, a constant (C) portion, and an antigen-binding domain, constituted by three CDR hypervariable regions. Four types of monomers (α, β, γ, and δ) are combined in αβ or γδ subunits to constitute functional forms of TCRs **(Figure 2B)**. The majority of circulating T-cells belongs to αβ type, with a specific homing to lymphoid tissues.

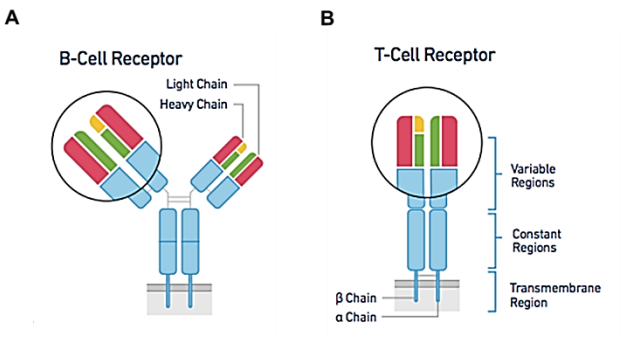

Figure 2. *A) Structure of Immunoglobulin or B-Cell Receptor. B) Structure of the T-Cell Receptor. Adapted from* ***https://www.10xgenomics.com/.***

Both in B- and T-cells, each receptor chain is encoded by several different gene segments, organized in gene complexes in different loci. Immunoglobulins are encoded by genes in three loci: immunoglobulin

heavy locus (IGH), mapping on chromosome 14, κ locus (IGK) on chromosome 2 and λ locus (IGL) on chromosome 22. Comparably, genes encoding for TCRs are organized in four loci (TCRA, TCRB, TCRG, and TCRD), each of them codifying respectively for a single α, β, γ or δ chain, and mapping on chromosome 14 for α and δ loci or on chromosome 7 for β e γ loci.

Hematopoietic stem cells and immature lymphoid progenitors display a germline configuration of Ig/TCR gene complexes, thus genes codifying for B- and T-cell receptors are in a nonfunctional status. In this configuration, each multigene superfamily contains one or few gene segments codifying for constant (C) regions and hundreds of gene segments codifying for variable regions. In particular, in the CDR3 portion of Ig/TCR gene complexes have been identified variable (V), joining (J) and diversity (D) gene segments, separated each other by non-coding regions. Diversity gene segments are present only in Ig heavy chain loci (IGH) and in those codifying for TCRs β and δ chain (TCRB and TCRD), while V, J, and C genes are common to all receptors loci. The expression of the antigen receptor starts with somatic recombination (rearrangement) of segments codifying for variable regions of receptors **(Figure 3)**. As a matter of fact, during the early B- and T-cell maturation, the pool of gene segments for each type of chain is randomly combined in different lymphocyte clones, thus each lymphocyte obtaining a peculiar combination of V-D-J gene segments.

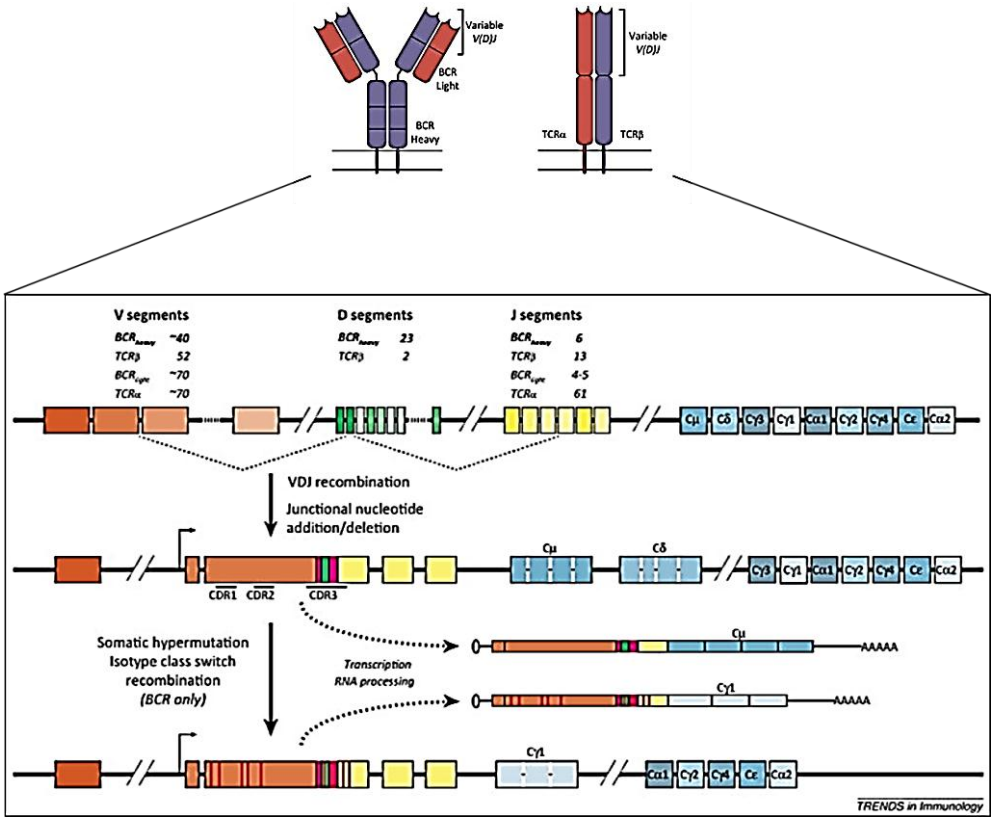

Figure 3. *Similar gene organization for Immunoglobulins (BCRs) and T-Cell receptors (TCRs). The antigen-binding surface is formed by the variable region of each chain, which is encoded by recombined V, J, and D segments. Antigen receptor diversification is generated during the lymphocyte maturation: for each B- or T-cell, V (orange), D (green), and J (yellow) gene segments are rearranged by somatic V(D)J recombination. During the rearrangement process, random insertion or deletion of nucleotides occurs at the junction regions (magenta), contributing to additional sequence diversity. Adapted from Calis and Rosenberg, 2014.*[32]

The V–D–J rearrangements are mediated by a recombinase enzyme complex, in which RAG1 and RAG2 proteins specifically recognize DNA recombination signal sequences (RSS), placed downstream of V gene segments, at both sides of the D gene segments and upstream of the J gene segments. Successively, RAG1 and RAG2 cut the DNA strand, determining the formation of palindromic sequences (P nucleotides). The first recombination event occurs between one D and one J gene segment. Then, this D-J recombination is followed by the joining of one V gene segment, from a region upstream of the newly rearranged DJ complex, forming a V-D-J gene segment. This process is responsible for the generation of receptor combinatorial diversity, which depends on the number of available V-D-J combinations. Combinatorial diversity is enhanced by random insertion and deletion of nucleotides at the junctional regions between rearranged gene segments, determining junctional diversity. This process occurs by the DNA repair enzymes through two different mechanisms. On one hand, during Ig/TCR gene rearrangement exonuclease enzymes delete from V, D or J gene segments one or more nucleotides. On the other hand, terminal deoxynucleotidyl transferase enzyme inserts random nucleotides even in the absence of the DNA template strand. It is not completely clarified whether the addition and deletion of nucleotides at the ends of coding regions occur in a defined order or simultaneously. Because of the junctional diversity process, junctional regions (N-regions) of rearranged immunoglobulin and T-cell receptor genes shall ensure that each Ig/TCR is different.

The Ig/TCR gene rearrangement process generally follows a hierarchical order during early lymphoid differentiation.[33] Concerning the B-cell maturation, firstly the IGH genes rearrange **(Figure 4)**. The protein produced by a functional rearrangement occurring on one allele exerts a negative feedback signaling, preventing the start of the rearrangement process on the other allele, according to allelic exclusion mechanism. Successively, somatic recombination occurs involving light chains loci κ, firstly, and λ secondly. During these events too, the κ chain expression determines a silencing signal direct against the second allele. Whether IGK rearrangement is unproductive on both alleles, the IGL genes rearrange. Finally, the newly formed functional light chains and heavy chains are combined to form the surface receptor.

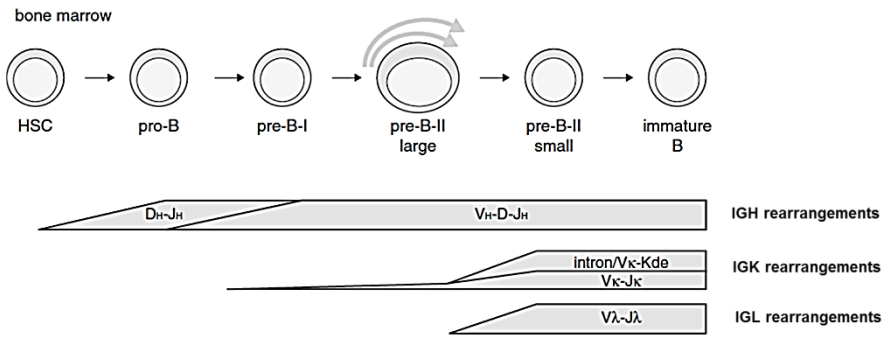

Figure 4. *Hierarchical order of Immunoglobulin rearrangements during B-cell differentiation in the bone marrow. Adapted from Langerak and van Dongen, 2012.*[33]

Comparably, during T-cells ontogenesis, pro-thymocytes, of whom genes are in germline configuration, once migrated in thymus undergo gene rearrangements starting from TCRD, TCRG and TCRB loci **(Figure 5)**. Rearrangements involving TCRA locus are late events since this gene region hosts the complete inner TCRD locus and can rearrange only in case of TCRD deletion caused by unproductive rearrangements. Therefore, in presence of productive γδ recombinations, the involved lymphocyte expresses this type of T-cell receptor, on the contrary, after allele exclusion the rearrangement process continues in TCRB locus and finally in TCRA, generating αβ subunit.

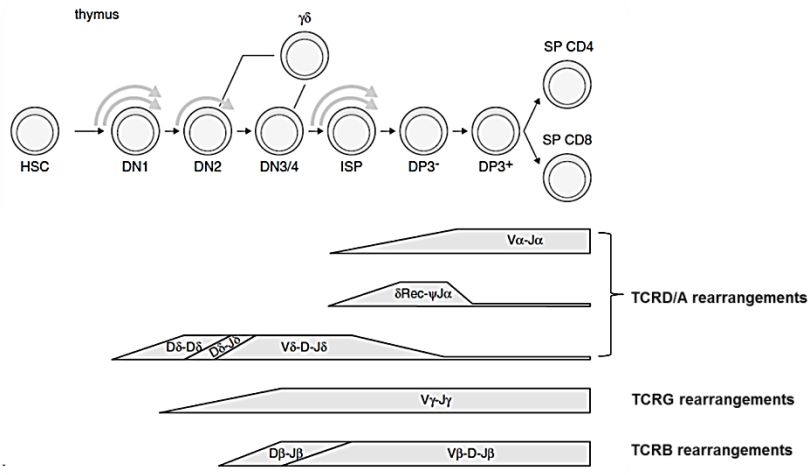

Figure 5. *Hierarchical order of T-cell Receptor rearrangements during T-cell differentiation in the thymus. Adapted from Langerak and van Dongen, 2012.*[33]

2.2.3. *Immunoglobulin/T-Cell Receptor rearrangements as clonal markers in Acute Lymphoblastic Leukemia for Minimal Residual Disease evaluation*

Immunoglobulin and T-cell receptor rearrangements assessment has a pivotal role in ALL since they are used as molecular markers for the diagnostic and prognostic evaluation. As a matter of fact, the random insertion and deletion of nucleotides at junctional regions of rearranged VDJ gene segments make the clone-specific junction regions of Ig/TCR genes as fingerprint-like sequences usable as leukemia targets. Unlike fusion genes, Ig/TCR gene rearrangements are physiological events not directly involved in ALL pathogenesis. However, since leukemia cells originate from a single transformed lymphoid precursor, all neoplastic lymphoblasts are supposed to contain the same clonal Ig/TCR rearranged gene, useful to detect a low number of ALL cells among a large number of normal lymphoid cells expressing gene rearrangements with different sequences.[34]

Ig/TCR rearrangements are not lineage-specific, then both B- and T-lineage leukemic cells can display cross-lineage rearrangements. Therefore, it is possible to identify immunoglobulin rearrangements in T-lineage ALL cases and, vice versa, T-cell receptor clonal markers in B-ALL cases. Nevertheless,

compared to T-ALL, rearrangements involving IGH, IGK, and IGL loci are more common in B-lineage patients, representing >95%, 30% and 20% of cases, respectively. On the contrary, TCRβ, γ and δ rearrangements are identified in 35%, 55% and 90% of B-lineage cases, respectively. T-lineage ALLs are characterized by the presence of TCR rearrangements in about 95-100% of cases, but in 20% of cases recombination involving Ig loci is assessed as well.

The MRD evaluation is based on the identification of clone specific junctional region, localized on CDR3 hypervariable region, and obtain complementary Allele-Specific Oligonucleotides (ASO)-primers. The ASO-qPCR approach combines the use of ASO forward primers with fluorescently labeled probes and reverse primers recognizing constant regions of the identified rearrangement. The ASO assay sensitivity is tested on ten-fold serial dilutions of DNA obtained from leukemic cells isolated at diagnosis with DNA from a pool of six-eight healthy donors. The MRD level is expressed as the logarithmic reduction of the leukemic clone detected at diagnosis **(Figure 6)**.

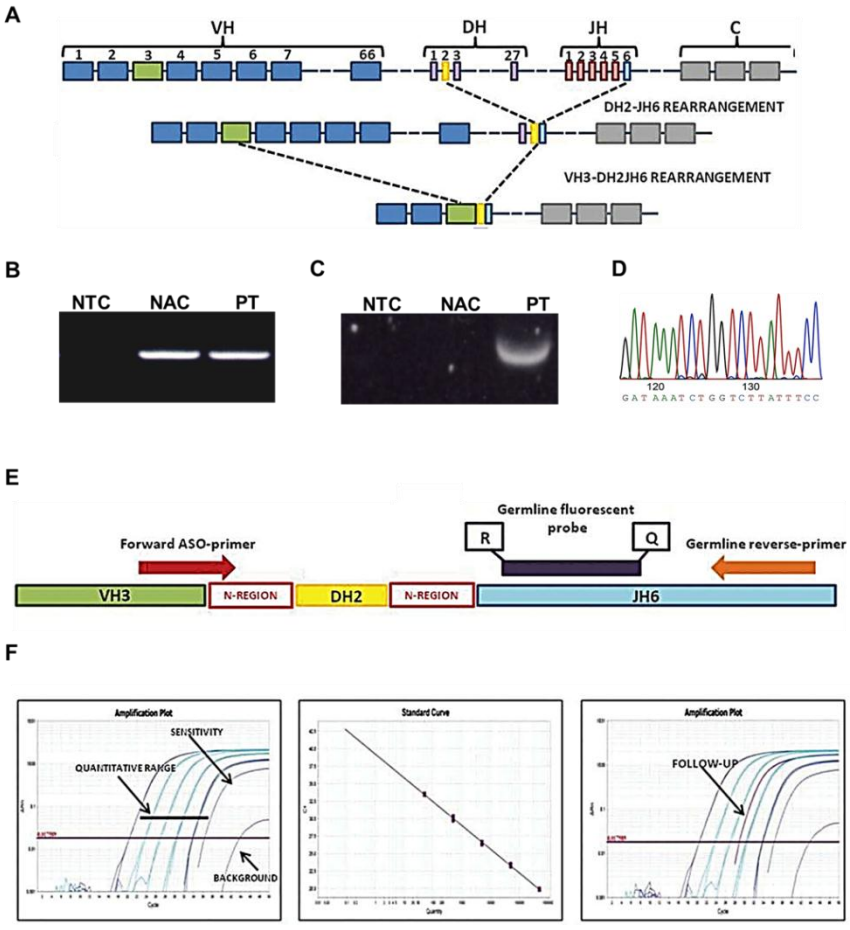

Figure 6. *Workflow for the MRD assessment. A) Ig/ TCR genes include discontinuous V, D and J segments, which undergo rearrangements during lymphocytes maturation. B) The identification of Ig/TCR rearrangements is performed at diagnosis by PCR amplification of the V-D-J segments involved in the Ig/TCR gene rearrangements C) Homo-heteroduplex clonality assay allows the discrimination of clonal rearrangements as clonality markers. D) V-D-J junction regions (N-regions), unique for each lymphocyte rearrangement, are defined by Sanger sequencing. E) The obtained sequences are used to design Allele-Specific Oligonucleotide (ASO) primers on N-junction regions, thus employed for MRD quantification by Real Time-qPCR in combination with primers and probes recognizing constant regions of the identified rearrangement. F) The MRD level is expressed as the logarithmic reduction of the leukemic clone detected at diagnosis using a patient-specific standard curve made by 10-fold dilutions. Adapted from Nunes et al., 2017.*[34]

Despite the advantage of the approach high sensitivity and patient-specificity, this procedure requires time, extensive knowledge and long-standing specialized personnel.[14] Moreover, its applicability is restricted to patients with Ig/TCR gene rearrangements with junctional regions suited to reach sufficient sensitivity. Another limitation of the technique is the inadequate or absent capability to identify the presence of oligoclonality at diagnosis or to monitor the clonal evolution of the leukemic clone, that can generate potential false-negative MRD evaluation during the follow-up of the patients. Indeed, monitoring MRD by ASO-qPCR requires to choose one or two representative clonal rearrangements for follow-up measurements, necessarily losing track of eventual minor sub-clones that might resist to therapy and be responsible for disease resistance or reappearance. Moreover, the limited number of oligonucleotides usable for a sample-consumption and economically-sustainable amplification panel does not allow the solution of rare rearrangements identification. Of note, another disadvantage of the ASO-qPCR approach is represented by the need of a large amount of DNA sample at diagnosis, required to perform relative quantification of the eventual residual disease burden, compared to disease onset.

Overall, in about 5-10% of translocations-negative ALL cases, the identification of suitable markers for MRD detection fails[30] and patients could not benefit of MRD-based treatment categorization due to this lack of suitable molecular probe.

2.3. *NEXT GENERATION SEQUENCING: The end of the Sanger era?*

Sequencing technology allows determining the order of nucleotides in DNA or RNA molecules, becoming one of the most powerful and fundamental tools in molecular biology. Since the importance and the wide applicability of sequencing, this technology improved more and more since its first application, assuring good throughput, accuracy, and sensitivity.

Sanger technology is the most widely-used method of DNA sequencing, based on the selective incorporation of chain-terminating dideoxynucleotides (ddNTPs) by DNA polymerase during *in vitro* DNA replication.[35] The incorporation of fluorescently labeled ddNTPs determines the stochastic termination of each round of primer extension. The label on the terminating ddNTP of any given fragment corresponds to the nucleotide identity of its terminal position, thus the complete sequence is resolved through electrophoretic separation in a capillary-based polymer gel. DNA fragments of discreet lengths exit the capillary and fluorescent labels are excited by a laser, coupled to four-color detection of emission spectra, providing the Sanger sequencing electropherogram.[36] Although it is possible to perform simultaneous electrophoresis in independent capillaries, one of the major limitations of the Sanger approach relies on the restricted level of sequencing parallelization.

On the contrary, Next Generation Sequencing (NGS) approach refers to innovative, massive parallel high-throughput sequencing technologies.[37] Indeed, a number of NGS platforms and approaches have been described in the last decade.[38,39] Despite the different used technologies, all platforms share the massive simultaneous sequencing of millions of DNA strands, allowing an increase in throughput yielding and a consistent reduction in terms of timing and costs of sequencing. Nevertheless, despite its superior performance, NGS is not yet routinely translated in clinical practice, rather conventional Sanger actually remains the most widespread, feasible technology.[1]

NGS technology can be applied to sequence the entire genome or specific areas of interest, such as coding regions (whole-exome sequencing) or small numbers of individual genes (targeted-sequencing). Targeted sequencing panels are focused on a select set of genes or gene regions of interest.[41] Two methods are commonly used for such targeted approaches: amplicon-based sequencing and capture hybridization-based sequencing.[41,42] Amplicons generation consists of the selected amplification of the targeted regions, lately purified by an oligonucleotides pool **(Figure 7A)**. By contrast, the capture-based approach requires the hybridization of the regions of interest with biotin-conjugated probes, thus recovered by streptavidin-coated magnetic beads **(Figure 7B)**. Therefore, the isolated gene regions are amplified by PCR. The amplicon-based workflow is simpler compared to the capture-based approach and it allows to cost-effectively amplify small portions of the genome, as well to efficiently identify small gene alterations such as Single Nucleotide Variants (SNVs) or insertions and deletions (indels). Conversely, the capture-based approach is the method of choice for the investigation of large regions of interest or a large number of genes. Furthermore, despite this

approach is a more laborious and time-consuming procedure, it allows to better investigate the mutational status of genes characterized by the presence of multiple and heterogeneous structural or numerical abnormalities. However, each technology amplifies single strands of a fragments library and performs sequencing reactions on the amplified strands.

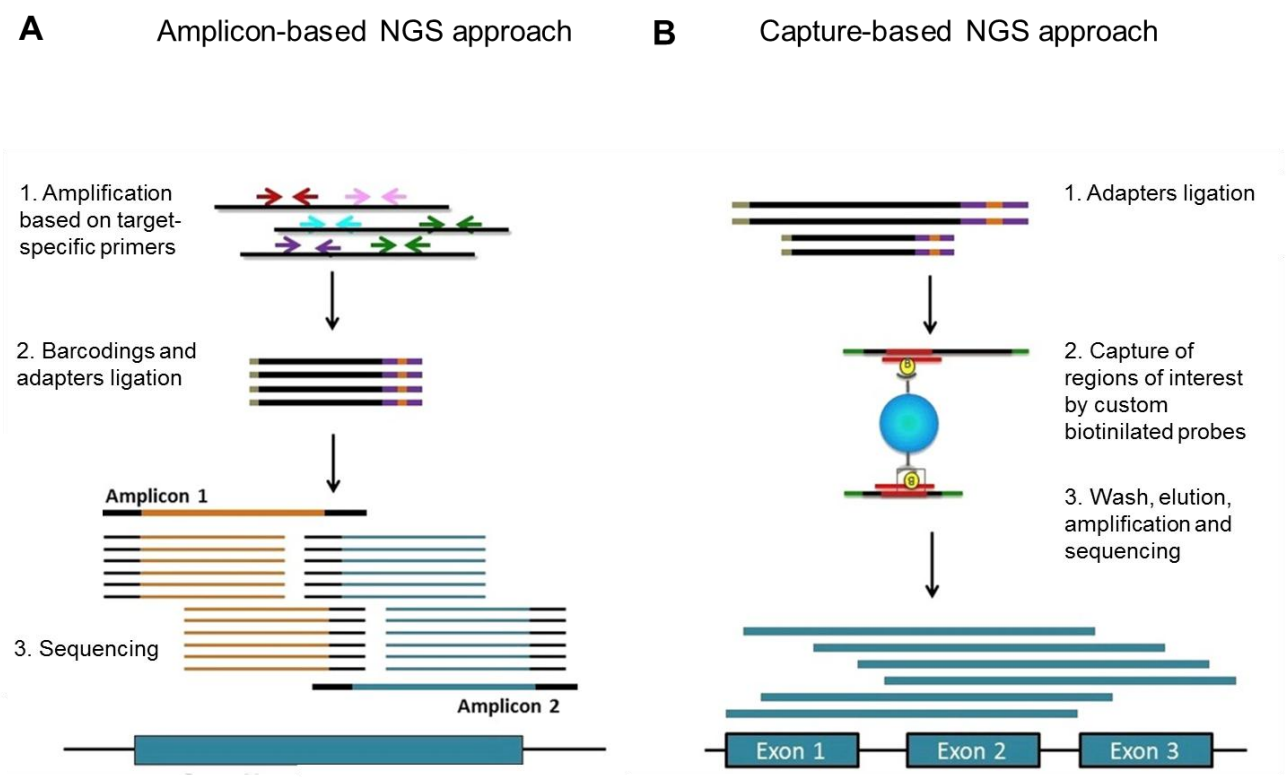

Figure 7. *A) Amplicon-based NGS simplified workflow. B) Capture-based NGS simplified workflow. Adapted from Jenning et al., 2017.* [43]

NGS is a multistep approach requiring library preparation, amplification reaction, sequencing, and data analysis.[38]

2.3.1. *Library preparation*

The first phase of the NGS workflow consists of library preparation. This process includes the amplification of regions of interest by PCR (amplicon-based NGS) or alternatively, the random genomic DNA (gDNA) fragmentation by enzymatic or mechanical approach, followed by the recovery of targeted regions through the hybridization to target-specific biotinylated probes (capture-based approach). However, both these procedures are preceded by the ligation of adapter sequences to PCR products or DNA fragments, to allow the later use of universal primers. Adapters are defined and

unique sequences used to cap the ends of each fragmented DNA molecule, which allow the hybridization to a solid surface, provide a priming site for both amplification and sequencing primers and finally act as barcoding (indexes) for multiplexing different samples in the same sequencing run **(Figure 8)**. As a matter of fact, differently from the traditional DNA sequencing approach, NGS selectively amplifies molecules by PCR thanks to adapter sequences. In single-end protocols, DNA molecules are cleaved into >1 kb size fragments and adapter sequences are ligated at one end of each DNA strand. On the contrary, in paired-end approach adapters are ligated at both ends of DNA fragment, allowing successively the two-direction sequencing of the same DNA molecule.[38,39]

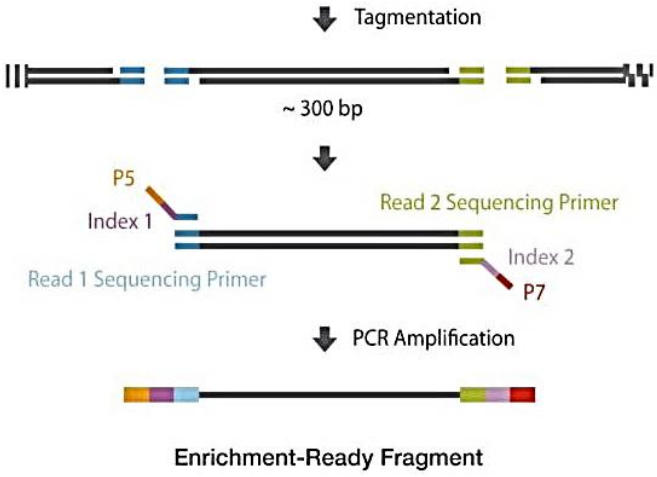

Figure 8. *Adapters ligation to PCR products or DNA fragments. Adapters oligonucleotides allow the hybridization of DNA fragments to universal primers for the subsequent library amplification and sequencing phases. Furthermore, adapters allow indexing specifically DNA fragments or amplicons. Adapted from* ***https://www.illumina.com***

2.3.2. *Library amplification reaction*

The DNA template amplification phase aims to increase the signal intensity for nucleotide detection. This phase is performed by emulsion PCR (emPCR) or by bridge amplification on solid-state. In emulsion PCR single-stranded DNA (ssDNA) hybridizes onto oligonucleotide bound beads. The beads are part of water-in-oil microemulsions containing all the components for PCR, acting as micro-reactor chambers. After the emulsion PCR amplification, thousands of complementary DNA strands are covalently bound to the beads, which are lately purified for the sequencing phase. On the contrary, amplification on a solid-state is performed by bridge PCR, using a solid surface coated with forward and reverse primers **(Figure 9)**. In this case, ssDNA fragments anneal randomly to the surface of the glass disposable flow cell, based on complementarity with the primers. The attached strand flips and bonds the other complementary primer. This phase is followed by the extension phase, leading to the

formation of a complementary strand. Finally, the two newly-formed strands are separated by denaturation and template DNA is washed. At the end of the process consisting of alternative extension and denaturation cycles, clusters of thousands of clonal DNA molecules are produced all over the solid surface starting from each single DNA template.

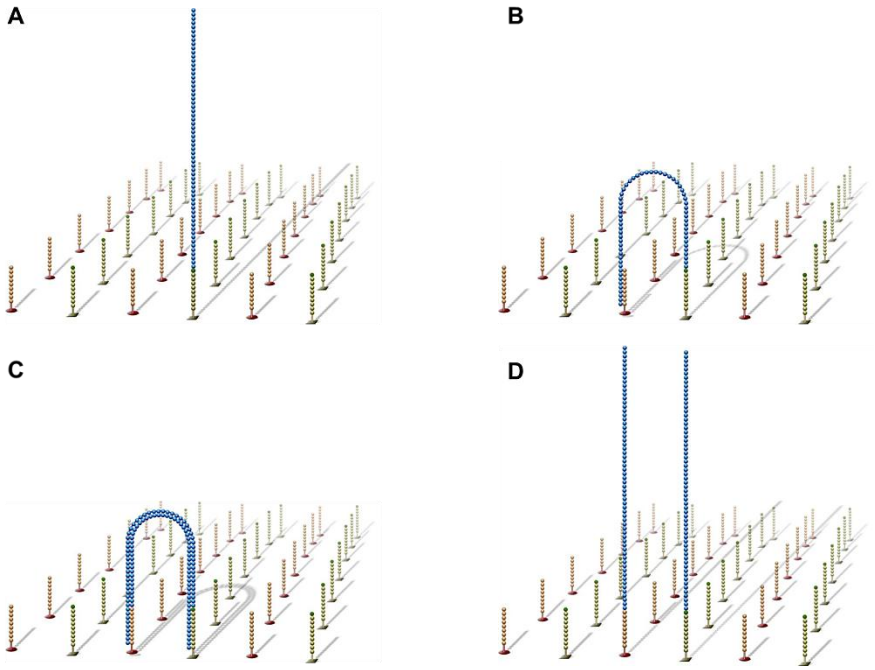

Figure 9. *Principle of bridge-PCR occurring on solid-state. A) ssDNA fragments, in which adapters have been ligated, bonds the complementary primer (in green) attached to the flow cell. B) The attached fragment flips and its free-end bonds the other complementary primer (in orange), generating a bridge. C) DNA Polymerase synthesizes a new strand starting from the bridge ssDNA template. D) The dsDNA molecules are separated by denaturation. Adapted from* ***https://www.illumina.com***

2.3.3. *Sequencing*

The sequencing methods employ Sequencing-by-Synthesis (SBS) or real-time Sequencing-by-Ligation (SBL) chemistry. The SBS approach includes Single-Nucleotide Addition (SNA) as well as Cyclic Reversible Termination (CRT) methods. SNA technology is based on the one-by-one incorporation for each cycle of a single dNTP into an elongation strand and the detection on newly-incorporated nucleotides can be performed by pyrosequencing or based on pH changes. The pyrosequencing-based Roche 454 technology works measuring the release of inorganic pyrophosphate during the nucleotide incorporation, which determines an enzymatic cascade responsible for the emission of a bioluminescent signal. In Ion Torrent technology, instead, the dNTP addiction is associated with a release of H^+ ions, determining a pH change detected by a semiconductor. Conversely, the Illumina technology employs DNA sequencing-by-synthesis technology with reversible terminator chemistry **(Figure 10)**. According to Illumina technology, nucleotides are conjugated with four different

fluorescent dyes and with a terminator group, and they are added in each sequencing cycle. Indeed, only one dye-terminated nucleotide is incorporated into the complementary strand because its conjugated terminator group prevents the attachment of the next nucleotide to the template DNA strand. For each nucleotide incorporation, the identification of the base is performed based on emitted fluorescence. After the removal of remaining unbound nucleotides and the cleavage of the dye and the terminator, a new nucleotide can be incorporated by DNA polymerase in the subsequent cycle. The single base calling is performed in parallel in each generated cluster.

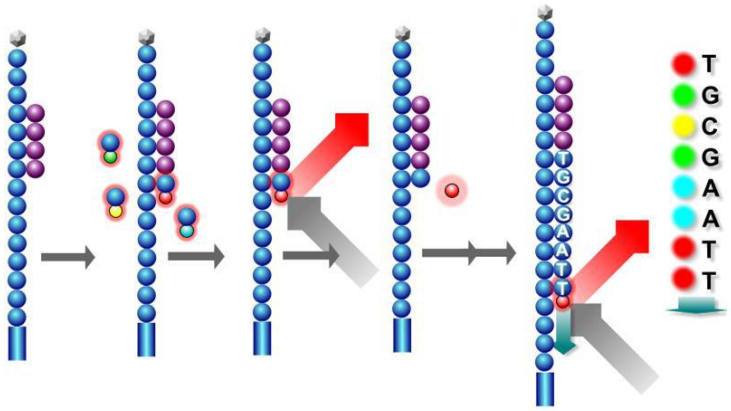

Figure 10. *Sequencing by Synthesis (SBS)-Cyclic Reversible Termination (CRT) technology. Adapted from* ***https://www.illumina.com***

Differently to Sequencing-by-Synthesis, the Sequencing-by-Ligation technology uses short 8-9 bp segments of DNA called oligonucleotides, rather than single bases. Oligonucleotides are fluorescently labeled probes, partially degenerated, because only the first two positions are occupied by known nucleotides. SBL chemistry is used by the SOLiD system, in which beads bounded to emulsion PCR products are fixed to a glass slide, hosting also a pool of labeled dinucleotide probes, ligases, and universal sequencing primers. After the probe ligation to DNA template by DNA ligase, the DNA sequence is determined by detection of a fluorescent signal, specific for the first of the two bases of the probes, in several cycles of DNA ligation and cycles of primer reset. For each next cycle, the newly synthesized DNA strand is denatured and hybridized by a new sequencing staggered primer, complementary to the n-1 position. In this technology, each nucleotide in the template is interrogated twice by two fluorescent signals, reducing the error rate.[38,39]

	Template DNA amplification	**Sequence**	**Detection**
Roche 454	emPCR	SBS-SNA	Emitted light
Illumina	Solid state	SBS-CRT	Emitted light
Ion Torrent	emPCR	SBS-SNA	Change in pH
SOLiD	Not Done	SBL	Emitted light

Table 4. *Summary of the different Next Generation Sequencing technologies so far described.*

2.3.4. *Bioinformatics and NGS data analysis*

NGS data analysis consists of three different steps, defined as primary analysis (base calling), secondary analysis (alignment and variant calling) and tertiary analysis. Primary analysis relies on the signal evaluation and the quality scoring of the base-calling. Secondary analysis consists of alignment on the reference genome (or the *de novo* assembly) and variant calling, while tertiary analysis requires the variant interpretation phase and eventually the candidate's validation and clinical decision-making process.

The primary analysis allows performing the base pairs calling starting from the signals (images or pH variations) detected during the sequencing phase and to determine the quality scores. The raw data file resulting from base calling is defined FASTQ, consisting in the combination of a string of A, T, C and G characters ("read"), with the relative quality score **(Figure 11)**.[44] The quality of the sequences is based on Phred score,[45] in which quality score (Q) is logarithmically related to the probability of base-calling error (e), according the following formula:

$$Q= -10\log10(e)$$

A Q score of 10, 20 and 30 corresponds to the probability to identify an incorrect base of 0.1 (10%), 0.01 (1%) and 0.001 (0.1%), respectively. Consequently, for a base with a Q30 score the probability that the sequenced base is incorrect is 1 in 1000 **(Table 5)**.

Quality Score	Probability of incorrect base call	Base call accuracy
Q10	1 in 10	90%
Q20	1 in 100	99%
Q30	1 in 1000	99.9%
Q40	1 in 10000	99.99%
Q50	1 in 100000	99.999%

Table 5. *Phred quality scores (Q) and relative logarithmically linked error probabilities and base call accuracies.*

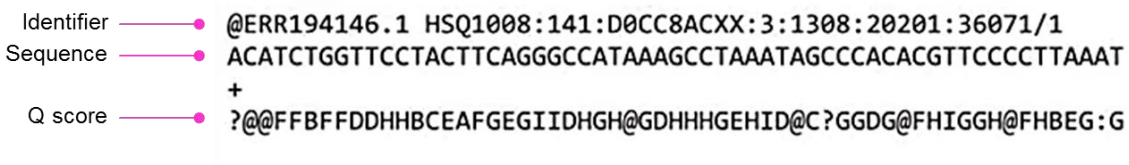

Figure 11. *Example of a FASTQ file format. The first line refers to the identifier of the generated sequence, the second line refers to the raw nucleotides sequence and the last line refers to the compressed version of the quality value obtained for each base. The compression of the Q score allows obtaining the same number of symbols as letters in the sequence, reducing storage space needed and processing costs. A separator, in this case a plus (+) sign, can be present as well between the sequence and the associated Q score.*

The secondary analysis relies on the alignment of the obtained reads to a reference sequence (or alternatively to the *de novo* assembly) and variant calling. Of note, this essential phase in NGS data analysis includes also the de-multiplexing step, by which the generated sequences are associated with the belonging samples thanks to indexes. Secondary analysis is performed by several algorithms and the alignment tools produce output data files in Binary Alignment/Map (BAM) format.[46] For each of aligned read, a Phred score is calculated, therefore the alignment output is associated with the probability of false alignment per base. This parameter can be modulated by the users, offering the possibility to allow more mismatch to the detriment of the alignment confidence. The number of reads covering a given genome locus is defined as depth coverage,[37] representing a data redundancy measurement and corresponding to the estimated result confidence. Generally, depth lower thresholds are well established, allowing the removal of inaccurate mismatches caused by sequencing errors and thus supported by very few reads. As well, the upper threshold setting permit to limit mismatches caused by PCR duplicates introduced by library preparation, copy number variations or caused by reads mapping on paralogous sequences. Importantly, the secondary analysis includes the variant calling, such as the identification of variations between a read and the reference genome, consisting of Single Nucleotide Variants (SNVs), insertions or deletions (Indels), large structural variants as translocations and Copy Number Variations (CNVs).

Tertiary analysis includes the evaluation of variants deleteriousness and phenotypical or clinical relevance, defined as annotation.[47] As a matter of fact, among thousands of filtered variants, only a percentage of them determines a biological effect, resulting in amino acids changes, in the untimely truncation of coding sequences or in the alteration of splicing sites or gene regulatory elements. Therefore, the annotation phase aims to separate neutral, non-functional variants from biological, disease-causing ones. For this purpose, several public databases are available and useful in combination with the comparison of relative variants frequency in samples and the common population. Indeed, a rare or absent variant in healthy control population is classically associated with pathogenicity, although this parameter is not sufficient to predict biological or clinical deleteriousness.

3. AIM

In Acute Lymphoblastic Leukemia (ALL) the evaluation of residual leukemia cells during and after treatment (Minimal Residual Disease, MRD) is considered the strongest predictive factor for relapse-risk assignment, making the MRD monitoring the standard of care in routine clinical practice.[21]

In about 30% of ALL patients, the molecular MRD assessment is based on the presence of recurrent chromosomal translocations with standardized, universal assays.[30,48] In the remaining 70% of patients, the availability of MRD data relies on surrogate markers: the identification of leukemia-specific Immunoglobulin (Ig) or T-Cell Receptor (TCR) clonal rearrangements. Although the diagnostic procedure is well-standardized within cooperative international working groups (EuroClonality and EuroMRD, ESHLO Foundation),[23] it requires the development of patient-specific reagents and assay conditions, making this technique laborious, time-consuming and demanding extensive expertise. Furthermore, in about 5-10% of cases this approach fails to identify suitable molecular markers for MRD assessment, thus resulting in the inability to allocate MRD-unknown patients in the adequate risk-category class and to offer them the most suitable therapy, avoiding under- or overtreatment.[30]

This failure can be attributed to techniques sensitivity issues as well as limits in PCR design. Recently, amplicon-based NGS methods have been described to identify clonal rearrangements at diagnosis and to monitor MRD in lymphoid malignancies.[49–51] Although highly promising, this approach is similarly based on selected amplification of most common Ig/TCR rearrangements,[52–57] therefore it may suffer from the same limitations as standard assays, that in a proportion of patients fail.[58,59]

The aim of this project was to develop and apply a capture-based NGS assay to better identify the Ig/TCR clonal rearrangements for the MRD classification of adult ALL patients. Furthermore, we enlarge the design of the NGS panel including several genes of interest, possibly useful for prognostication or MRD detection in ALL.

For this purpose,

- we designed a capture-based NGS panel to bind and capture the Ig and TCR chromosomal regions and, subsequently, some genes of interest in ALL;
- we validated it on samples from patients well characterized for clonality by conventional method[28,60] within the NILG-ALL 09/00 clinical trial.[15,16] We also evaluated different bioinformatics tools for data analysis and rearrangements identification;
- we applied our capture-NGS panel to patients lacking molecular probes (MRD-unknown) within the NILG-ALL 09/00 clinical trial and we performed MRD evaluation, reclassifying them as high risk/low risk based on the new NGS findings;
- finally, we verified the prospective feasibility of the capture NGS approach to the MRD-unknown cohort of patients within the prospective national treatment protocol for adult ALL therapy.

4. MATERIALS AND METHODS

4.1. *Clinical samples*

For the validation phase of the study, we tested diagnostic bone marrow material from 10 adult ALL patients (6 B- and 4 T-lineage, median age 33 years, 3 females and 7 males), enrolled into the NILG-ALL 09/00 clinical trial (ClinicalTrials.gov identifier: NCT00358072)[15,16] and previously studied for clonality assessment following the EuroMRD guidelines.[28,60] Patients have been selected for having at least 50% of blasts in the diagnostic specimen, two sensitive patient-specific probes and for being MRD positive (n=5) or MRD negative (n=5) at the end of consolidation phase of treatment **(Table 6)**.

We also included Human Umbilical Vein Endothelial Cells (HUVECs) and mesenchymal cord blood cells, both tested in duplicate, as negative controls for Ig/TCR rearrangements by reason of their germline configuration of Ig/TCR loci, in order to verify the absence of false-positive results during the preparation of the libraries, sequencing or data analysis phase.

Patient ID	Age at DX	Sex	ALL lineage	Blasts %	Karyotype	Probe 1 (P1)	P1 assay Sensitivity	Probe 2 (P2)	P2 assay Sensitivity	MRD
BG_371	41	M	B	80	48,XY, +?8,+?13(9)/46,XY (12)	TRGV3*02-J1*02	10^{-5}	TRDV2*02-D3*01	10^{-4}	POS
BG_4502	28	M	T	100	46,XY del(6)(q21), t(9;10)	TRGV9*01-J1*02	10^{-5}	TRDV1*01-J1*01	10^{-5}	NEG
BG_5038	56	M	B	99	46,XY	TRDD2*01-DD3*01	10^{-5}	IGKV1-16*02-KDE	10^{-5}	POS
BG_5418	39	M	B	90	46,XY	IGHV3-7*02-J6*02	10^{-5}	TRGV9*01-J1*02	10^{-5}	POS
BG_5452	25	M	T	100	5q- (q31)	TRGV10*02-J1*02	10^{-4}	TRBD1-J2-3	10^{-4}	NEG
BG_9574	46	M	T	80	46,X,inv(Y),del(6)(q21),t(10;14)(q24;q11)(12)/46idem, add(9)(p?)(4)	TRGV11*01-J2*01	10^{-5}	TRGV4*02-J1*02	10^{-4}	POS
BG_9813	16	F	T	90	46,XX. Molecular detection of SIL-TAL1 fusion gene	TRDD2*01-J1*01	10^{-5}	TRGV9*01-J2*01	10^{-4}	NEG
BG_11360	15	F	B	60	(46,XX), 1t(1;12), 1del(9)	TRDD2*01-D3*01	10^{-4}	TRDV2*02-D3*01	10^{-4}	NEG
BG_11720	41	F	B	90	46,XX	TRGV9*01-J1*02	10^{-5}	IGHV6-1*01-J6*02	10^{-5}	NEG
BG_11806	27	M	B	90	46,XY	TRGV5*01-J1*02	10^{-4}	IGHV1-2*02-J5*02	10^{-5}	POS

Table 6. *Clinical and biological characteristics of adult ALL patients evaluated in the validation of the capture-NGS panel, formerly studied for clonality assessment following the conventional EuroMRD guidelines within the NILG-ALL 09/00 clinical trial. Patients have been selected for having at least two sensitive patient-specific probes and for being MRD positive (n=5) or negative (n=5) at the end of the consolidation phase.*

The validated capture-based NGS method was then applied to residual diagnostic bone marrow or peripheral blood samples from 23 adult ALL patients (13 B- and 10 T-lineage, median age 43 years, 8 females and 15 males) enrolled into the NILG-ALL 09/00 trial, in which the identification of a suitable MRD molecular marker failed or was particularly difficult (MRD-unknown). Also in these cases, the blasts percentage in the samples was >50% in all cases except one (BG_4005), for which the blasts percentage at disease onset was not available **(Table 7).**

Patient ID	Age at DX	Sex	ALL lineage	Blasts %	Karyotype
BG_855	63	F	B	>50	46,XX
BG_1125	43	M	B	>50	Not Done: *punctio sicca*
BG_2097	62	M	B	60	47,XY,inv dup(iq)(ipter 1q22;1q22cen;1q23)1qter),dup(2p)(2pter)2p13;2p13)qter),t(8;14)(8q24;14q32),del(13)(13pter)13q13;13q22)13qter),+mar,46,XY
BG_2481	32	M	T	>50	Absence of Metaphases
BG_3895	20	M	T	>50	46,XY
BG_4005 *	18	M	T	Na	46,XY
BG_4255	36	M	T	>50	46,XY
BG_4254	63	F	B	90	63-65,XX(7)/46,XX(2)
BG_4379	50	M	T	>50	46,XY (10)
BG_5702	62	F	B	85	Hyperdiploid (63 chr)
BG_6037	33	F	T	>50	Absence of Metaphases
BG_12438	26	M	T	>50	46,XY
BG_6490	63	M	B	95	Triploidy in 5 Metaphases/15 normal Metaphases
BG_8646	59	M	T	90	46,XY,del(5)(q22q35),del(12)(p11p13),del(17)(p11)(18); 46, XY (2)
BG_8345	33	F	B	64	46,XX
BG_9445	45	M	B	80	36,XY,-2,-3,-4,-7,-8,-9,-d,-16,-17,-20
BG_10112	63	M	B	93	46,XY, t(4;11)(q21;q23)
BG_10487	20	M	B	>50	Not Evaluable
BG_10640	62	F	B	>50	Absence of Metaphases
BG_11053	25	F	B	>50	46,XX(18). Molecular detection of KMT2A-AFF1 transcript from t(4;11)
BG_11269	33	M	T	>50	46,XY
BG_11345	63	M	T	>50	48,XY+10+21 / 47XY, 10/47,XY+21 / 46XY
BG_11584	38	F	B	100	Absence of Metaphases. Molecular detection of KMT2A-AFF1 transcript from t(4;11)

Table 7. *Clinical and biological characteristics of adult ALL patients enrolled into the NILG-ALL 09/00 clinical trial in which the identification of a suitable Ig/TCR molecular marker for MRD evaluation failed or was particularly difficult (MRD-unknown). * For patient BG_4005 the blasts percentage in the diagnostic sample was not available (Na).*

This cohort of patients had also undergone a previous, independent NGS analysis on *TP53* gene, that proved the very bad prognostic impact of its mutation on adult ALL patients.[61] Additional 29 ALL diagnostic samples, of which 4 cases proved *TP53*-mutated and 25 wild-type, were retrospectively evaluated for Ig incomplete rearrangements with conventional method to explore correlation between the presence of incomplete rearrangements and *TP53* mutations. Patients' clinical and biological characteristics are reported in **Table 8**.

Patient ID	Age at DX	Sex	ALL lineage	Blasts %	Karyotype	Genetics TP53
BG_4205	30	F	B	70	46,XX	MUT+DEL
BG_2873	37	M	B	70	Poor quality metaphases; MLL (11q23)	MUT
BG_11543	62	F	B	70	46,XX	MUT+DEL
BG_10442	44	M	T	40	54,XY,+5?,+7,+8,+8,+15?,+21?,+21,+22(6)/46,XY(1)	MUT
BG_8142	45	M	T	90	46,XY; 46,XY,t(11;?); 48,XY,+aberrations. Molecular detection of del1 type II. Poor quality metaphases and low mitotic index	MUT
BG_630	35	F	B	>50	46XX, +9p/46XX	WT
BG_2145	32	F	B	>50	46,XX	WT
BG_2195	30	F	B	>50	46,XX, idic(9)(p12or13), der(19)t(1;19)(q23;p13)	WT
BG_2567	22	M	B	>50	Na	WT
BG_4612	41	F	B	>50	46,XX(2)/45, XX,-9,der(13)t(9;13),der(16),der (22)(5)/45,XX,-9,der(13)t(9;13),der(16)(1)	WT
BG_4790	24	F	B	>50	46,XX	WT
BG_5077	40	F	B	>50	46,XX	WT
BG_5136	44	F	B	>50	Na	WT
BG_5956	26	M	B	>50	46,XY	WT
BG_6333	45	F	B	>50	47XX, +14 (10)	WT
BG_6380	24	F	B	>50	46,XX	WT
BG_6533	56	M	B	>50	Na	WT
BG_7275	48	M	B	>50	Na	WT
BG_7567	31	M	B	>50	Na	WT
BG_8206	47	M	B	>50	46, XY, -1,+der(1)t(1;?)(p36;?);44-46,XY	WT
BG_8632	32	F	B	>50	46,XX	WT
BG_8844	39	F	B	>50	46,XX	WT
BG_10981	21	M	B	>50	46, XY , del9	WT
BG_11360	15	F	B	>50	46XX, 1t(1;12), 1 del(9)	WT
BG_11568	21	F	B	>50	Na	WT
BG_11696	19	M	B	>50	Na	WT
BG_11784	33	M	B	>50	46,XY	WT
BG_11958	38	F	B	>50	46,XX	WT
BG_12374	23	F	B	>50	Na. Molecular detection of TCF3-HLF transcript from t(1;19)	WT
BG_1839	40	M	B	>50	46XY, 46XY +3q	WT

Table 8. *Clinical and biological characteristics of adult ALL patients enrolled into the NILG-ALL 09/00 clinical trial analyzed for the mutational status of the TP53 gene by an independent NGS approach. These patients have been retrospectively tested for the presence of incomplete IGH clonal rearrangements by conventional methods. *For patients indicated with *, the karyotype evaluation was not available (Na).*

The enhanced version of the NGS panel (v2), targeting both Ig/TCR loci and 25 ALL-related genes, was verified on diagnostic bone marrow samples from 7 adult ALL patients within the NILG-ALL 09/00 trial, formerly evaluated for Ig/TCR clonal rearrangements by standard procedure and/or by capture-NGS approach (BG_5038, BG_10442, BG_21292, BG_4255, BG_4379, BG_11269 and BG_11584). Also, this capture-NGS panel was then applied for the analysis of 12 new diagnosed ALL patients MRD-unknown (3 B-ALL, 7 T-ALL and 2 T-Lymphoblastic Lymphoma (T-LL), median age 40 years, 1 female and 11 males). Seven out of 12 patients were enrolled within the active GIMEMA (Gruppo Italiano Malattie EMatologiche dell'Adulto) LAL1913 clinical trial for adult ALL treatment (ClinicalTrials.gov

identifier: NCT02067143). The studied cases had a blasts percentage in the diagnostic sample ranging from 5% (BG_37265 and BG_39541) to up 50%. The patients 'clinical and biological characteristics are listed in **Table 9**.

Patient ID	Age at DX	Sex	ALL lineage	Blasts %	Karyotype
BG_37265	52	M	T-LL	5-10	46,XY (15). Poor quality metaphases; it is not possible to exclude the presence of structural aberrations.
BG_41985*	21	M	B	95	Na
BG_41182*	50	M	B	95	Na
BG_41209*	20	M	T	63	Na
BG_41733*	21	M	T	70	Na
BG_41165*	63	M	T	71	Na
BG_41408*	36	M	T	70	Na
BG_40129*	43	M	T	96	Na
BG_39541*	28	M	T	5	Na
BG_39652*	35	M	T-LL	30	Na
BG_42228*	44	M	T	83	Na
BG_42309	75	F	B	70	46,XX (4). Low number of metaphases; it is not possible to exclude the presence of other cell lines.

Table 9. *Clinical and biological characteristics of newly diagnosed adult ALL patients missing suitable MRD assay by standard approach (MRD-unknown) prospectively evaluated by capture-based NGS panel v2. *For patients indicated with *, the karyotype evaluation was not available (Na).*

4.2. *Biological samples processing and Mononuclear cells separation*

The separation of bone marrow/peripheral blood-derived Mononuclear Cells (MNCs) from other cell populations was performed by density gradient separation (Lympholyte® Cell Separation Media, Cedarlane Laboratories, Burlington, ON, Canada). Briefly, anticoagulant-treated samples were carefully layered over the Ficoll solution (ratio Ficoll/blood ranging from 1:1 to 1:2) and centrifuged at 2000 rpm for 20 minutes at room temperature (RT). The mononuclear cells stratum, at the interface between the plasma and the Ficoll layer, was transferred in a sterile centrifuge tube and washed twice with Phosphate Buffered Saline (PBS) 1X, followed by centrifugation at 2000 rpm for 5 minutes at RT. After the second wash, cells were gently resuspended in a known PBS volume and counted with Coulter ACT (Beckman Coulter, Brea, CA, USA). Then, after centrifugation, the supernatant was discarded by inversion and the dry pellet has been stored at -20°C in 1.5-2 ml tubes.

4.3. *DNA extraction and quantification*

The extraction of genomic DNA (gDNA) was performed by Gentra Puregene Cell Kit (Qiagen, Hilden, Germany) or by Maxwell® LEV DNA Purification Kit (Promega, Madison, WI, USA). According to the Gentra protocol, the dry cell pellet was resuspended in 1 ml of Cell Lysis solution and added with 300 µl of Protein Precipitation solution. The obtained solution was mixed by vortex and centrifuged at 13000 rpm for 3 minutes at RT. The supernatant containing DNA was transferred to a new tube containing an equal volume of isopropanol and mixed by inversion until the gDNA formed a white filamented aggregate. The tube was centrifuged 13000 rpm for 2 minutes at RT, the supernatant was discarded and the DNA pellet was washed with 300 µl of 70% ethanol and centrifuged again 13000 rpm for 2 minutes at RT. After the supernatant discard, the DNA pellet was let drying at 55°C a few minutes and then resuspended in the DNA Hydration Solution.

Alternatively, the gDNA extraction was automatically performed by Maxwell® instrument and dedicated kit, according to the producer recommendations. Each pellet sample was resuspended in 300 µl of PBS 1X, then 30 µl of Proteinase K solution and 300 µl of Lysis Buffer were added to each tube. The tubes were mixed by vortex for 30 seconds and incubated at 56°C for 20 minutes. During the incubation, in the Maxwell® 16 LEV Cartridge Rack a precast cartridge was prepared for each sample. Into the specific position of each cartridge one plunger and an empty Elution Tube, containing 50 µl of Elution Buffer, were placed. Once the incubation was completed, each blood lysate sample was transferred to the specific well of the cartridge and the Maxwell® 16 LEV Cartridge Rack containing the prepared cartridges was placed on the Maxwell® 16 Instrument platform for the automated purification of gDNA. The Maxwell® 16 Instrument purifies samples using a paramagnetic particle that utilizes cellulose-based binding of nucleic acids, which provides a mobile solid phase for capture, washing, and purification of DNA.

The qualitative and quantitative gDNA quantification was performed by spectrophotometric analysis using the NanoDrop system (NanoDrop ND-1000 Spectrophotometer V3.7, Thermo Scientific, Waltham, MA, USA). The DNA concentration is calculated based on the absorbance (A) at 260 nm, 280 nm, and 230 nm. The quality of the samples has been determined based on the ratios $A_{260/280}$ and $A_{260/230}$: samples were considered of good quality if absorbance ratios were 1.8-2.0 and 1.6-2.4 respectively, indicating DNA molecule integrity and lack of proteins contamination.

To guarantee the quantification accuracy, the gDNA concentration in the samples was also estimated by a fluorometric approach using the Qubit® 3.0 system (Thermo Scientific, Waltham, MA, USA). The Qubit® fluorometer utilizes fluorescent dyes, specific to the target of interest, that emit only when bound to the target molecules, even at low concentrations. This technology avoids the unspecific measurement of absorbance at 260 nm, thus preventing the overestimation of the DNA concentration.

4.4. *Capture-based NGS panel design*

We designed a custom capture-based NGS panel (EZ SeqCap gene panel, Roche NimbleGen, Madison, WI, USA), based on gene coordinates derived from the GRCh38/hg38 gene assembly targeting coding V, D and J genes in the Ig/TCR loci for the identification of V-D-J and D-J rearrangements (panel v1). When we enhanced the panel we included both Ig/TCR loci and enrichment probes targeting coding regions of a selected group of 25 genes relevant in a diagnostic or prognostic setup listed in **Table 10** (panel v2).[62] The first panel, containing only the V, D, J and constant regions of the Ig/TCR loci, incorporated about 180 kb, while the panel containing also the other relevant genes embraced 350 kb.

Gene name	Chromosome
ALB	chr4
BRAF	chr7
CREBBP	chr16
EZH2	chr7
FBXW7	chr4
FLT3	chr13
IDH1	chr2
IDH2	chr15
IKZF1	chr7
IL7R	chr5
JAK1	chr1
JAK2	chr9
JAK3	chr19
KRAS	chr12
NOTCH1	chr9
NRAS	chr1
PAX5	chr9
PTEN	chr10
SH2B3	chr12
STIL	chr1
TAL1	chr1
TAL2	chr9
TET2	chr4
TP53	chr17
TYK2	chr19

Table 10. *List of genes included in the enhanced version of the capture-NGS panel (panel v2).*

4.5. *Libraries preparation*

A gDNA amount ranging from 0.6 to1µg was diluted in Tris-HCl solution and subjected to fragmentation to an average size of 600 bp either by sonication or enzymatic digestion. For the sonication approach, 53 µl of gDNA solution were transferred in specific snap-cap Covaris tubes, then inserted twice for 50 seconds in the Covaris S220 Focused-ultrasonicator (Covaris, Woburn, MA, USA), following the producer recommendations to set the instrument parameters. For enzymatic fragmentation, 35 µl of gDNA containing solution were incubated with the fragmentation master mix using Kapa enzymes (Kapa Biosystems, Wilmington, MA, USA) for 18 minutes at 37°C and put on ice to block the fragmentation reaction.

Successively, libraries in the pre-hybridization step were prepared with the KAPA HyperPluys Kit (Kapa Biosystems, Wilmington, MA, USA) according to the manufacturer's instructions. Fragmented samples were incubated with the end-repair and A tailing master mix, mixed by vortex few seconds and incubated in the 2720 Thermal Cycler (Applied Biosystems, Foster City, CA, USA) for 45 minutes at 65°C. After the end-repair and A tailing reactions samples were supplemented whit 5 µl of SeqCap Library Adapter 10mM and with 45 µl of ligation master mix, then incubated for 30 minutes at 20°C. Once completed the ligation reaction, we performed the post-ligation cleanup, incubating the PCR products with 88 µl of AMPure XP beads (Beckman Coulter Inc., Brea, CA, USA) for 5 minutes at RT. AMPure XP magnetic beads ligate specifically DNA and, after the placing on a magnetic particle collector, allow to discard the supernatant containing the unbound molecules. AMPure XP beads have been washed twice with 80% ethanol for 30 seconds at RT, let drying at RT for the ethanol evaporation and resuspended with 53 µl of elution buffer (Tris-HCl 10 mM, pH 8.0). After the incubation of the beads for 2 minutes at RT for the elution step, tubes were placed in a magnetic particle collector and 50 µl of supernatant containing gDNA were transferred in new tubes. The gDNA fragments of interest, ranging from 250 bp to 750 bp, were selected by dual size selection, using different sequential concentrations of AMPure XP beads according to the manufacturer instructions.

Samples underwent the pre-capture PCR amplification, using KAPA Hifi HotStart ready mix and Library amplification primer mix in a total volume of 50 µl and incubated on the 2720 Thermal Cycler following the Pre-Capture LM PCR amplification thermal protocol:

Temperature	Time	Cycles
98°C	45"	1
98°C	15"	2
60°C	30"	
72°C	30"	
72°C	60"	1
4°C	∞	1

Pre-capture PCR products were cleaned up by AMPure XP beads as previously described and each amplified sample library was accurately quantified by NanoDrop, Qubit® and by Agilent Bioanalyzer DNA 1000 assay (Agilent Technologies Inc., Santa Clara, CA, USA), following the manufacturer instructions. The recommended values to pass the quality check were 1-7-2.0 for the $A_{260/280}$ ratio measured by NanoDrop, 1-2 µg for the sample library yield measured by Qubit® and 150-500 bp for the fragments size distributions (with a peak around 300-350 bp) estimated by Bioanalyzer.

Therefore, libraries were grouped in a single pool containing an equal mass of different libraries and reaching a combined DNA mass of at least 1.25 µg. In a new PCR tube, 5 µl of COT Human DNA (1 mg/ml) and 2 µl of the Multiplex Hybridization Enhancing Oligo Pool were added to 1 µg of libraries pool. After the determination of the mixture total volume, two volumes of AMPure XP beads were added to the tube and incubated for 10 minutes at RT. The tube was placed on the magnetic particle collector, the supernatant was discarded and beads were washed twice with 80% ethanol. After the ethanol removal, the hybridization master mix was incubated for 2 minutes at RT with the bead-bound pool sample. The tube containing the mixture was placed one more time on the magnetic particle collector and the supernatant was transferred in a new tube to being hybridized to our custom-designed EZ SeqCap gene panel (Roche NimbleGen, Madison, WI, USA) through an incubation in the thermocycler for 10 minutes at 95°C and for 16-20 hours at 47°C. After the hybridization, the recovery of the regions of interest required the resuspension of the hybridized pool with the streptavidin-conjugated capture beads and the incubation of the mixture for 45 minutes at 47°C. The product of the capture reaction underwent the cleanup with specific wash buffers to remove reagents and probes, in order to avoid the probes bond to unspecific regions and to reduce the off-target percentage.

The captured pool was then amplified by Middle-Capture LM PCR, using KAPA Hifi HotStart ready mix and Post-LM-oligos, according to the following thermal protocol:

Temperature	Time	Cycles
98°C	45”	1
98°C	15”	5
60°C	30”	
72°C	30”	
72°C	60”	1
4°C	∞	1

Middle-capture PCR products were successively purified using AMPure XP beads as previously described. To enhance the enrichment and the specificity of the captured regions, the 16-20 hours hybridization, the 45 minutes capture, the post-capture cleanup, and the amplification were repeated. Indeed, the pooled-library underwent the post-capture PCR amplification, using KAPA Hifi HotStart

ready mix and Post-LM-oligos and incubated in the 2720 Thermal Cycler following the Post-Capture LM PCR amplification thermal protocol:

Temperature	Time	Cycles
98°C	45”	1
98°C	15”	14
60°C	30”	
72°C	30”	
72°C	60”	1
4°C	∞	1

After the post-capture PCR amplification, the pool was purified by AMPure XP beads and quantified by NanoDrop and Qubit® systems, then the size of the DNA fragments was estimated by Agilent Bioanalyzer High Sensitivity technology as previously described. Based on the size and the measured concentration, the molarity of the pool was calculated and 12 pmol of the pool were denaturated by NaOH 0.1 N for 5 minutes at RT and diluted with HT1 reagent (Illumina, San Diego, CA, USA) to stop the denaturation. The 12 pM pool was loaded in a v3 cartridge for the 300 bp paired-end sequencing on a MiSeq platform (Illumina, San Diego, CA, USA).

4.6. *NGS data analysis*

Bioinformatics data were processed separately for Ig/TCR rearrangements and gene mutations. The identification of Ig/TCR rearrangements was performed through the open-source application Vidjil (www.vidjil.org),[63,64] loading for each sample the NGS raw-data FASTQ files R1 and R2. As a matter of fact, Vidjil-algorithm processes high-throughput sequencing data to extract V(D)J junctions and gather them into clones, detecting gene rearrangements from both Ig and TCR, as well as some incomplete or uncommon rearrangements.

Since no official guidelines for the clonality interpretation in the contest of NGS technology were so far drawn up, based on the published data,[65] we considered as clonal those rearrangements that were present at a level >5% in the same locus and contextually >1% relatively to the total amount of rearrangements in a given sample.

Variant calling in the capture included ALL relevant genes was performed by MiSeq® Reporter Software and visualization of results was possible by the Illumina VariantStudio Data Analysis Software (Illumina, San Diego, CA, USA) and the Integrative Genomics Viewer tool (http://software.broadinstitute.org/software/igv/).[66,67] Variants were considered only if reaching at least a read depth of 100 reads and a Variant Allelic Fraction (VAF) of 10%. The pathogenicity

prediction of variants was verified based on ClinVar public archive, reporting the association between human variations and phenotypes, with supporting evidence regarding their clinical significance, and based on the Catalogue of Somatic Mutations in Cancer (COSMIC), an expert manually curated database of somatic mutations in human cancers. Moreover, the frequency in the general population of the identified variants was verified through ExAC, GnomAD, and dbSNP databases, in order to exclude Single Nucleotide Polymorphisms (SNPs) and benign variants, with no biological or clinical impact.

4.7. *Clonality assessment validation*

In all the patients the Ig/TR gene rearrangements had been determined at diagnosis by polymerase chain reaction (PCR), heteroduplex clonality assay and Sanger sequencing according to BIOMED protocols.[28,60] The heteroduplex assay required the denaturation of the PCR products at 94°C for 5 minutes and the subsequent renaturation at 4°C for one hour, to induce the formation of DNA dimers, separated by 6% polyacrylamide gel electrophoresis at 40 mA. After the renaturation, homo-dimers deriving from monoclonal rearrangements were visualized as distinct bands in polyacrylamide gel, while polyclonal rearrangements produced hetero-dimers resulting in a smear. Also, in the case of an oligoclonal cell population, the presence of both homo- and heteroduplexes resulted in a variable number of bands.

Monoclonal PCR products underwent Sanger sequencing (GATC services by Eurofins GATC Biotech GmbH, Konstanz, Germany) and the Sanger electropherogram was analyzed by Sequencing Analysis Software v5.2 (Applied Biosystems™, Foster City, CA, USA). The finger-print CDR3 regions of D-J/V-D-J gene rearrangements were detailed through the IMGT V-Quest web application (www.IMGT.org).

Furthermore, additional rearrangements found by NGS, but not detailed in BIOMED protocols, were validated by custom-made PCR oligonucleotide design by Primer Express® Software v2.0 (Applied Biosystems™, Foster City, CA, USA) or under the need of a second re-amplification to enhance sensitivity. In case of a second re-amplification, polyacrylamide-derived bands, stored at +4°C after the clonality heteroduplex assay, were added with 380 µl of elution buffer (Tris-HCl 10 mM, pH 7.4; SDS 0.1%, EDTA 1 mM, pH 8.0) and incubated overnight at 55°C. The aqueous phase derived from each band was transferred in a new tube and added with 760 µl of cold absolute ethanol and 38 µl of sodium acetate 3M, pH 5.2. Tubes were mixed by inversion and incubated at least 30 minutes at -20°C, then centrifuged at 13000 rpm for 20 minutes at +4°C. After the supernatant removal, the DNA forming a small pellet was washed by 500 µl of cold 70% ethanol and once again centrifuged at 13000 rpm for 5 minutes at +4°C. The supernatant was discarded, the pellet was let drying 10 minutes under a fume hood and re-hydrated with 30 µl of nuclease-free water. According to PCR protocols based on the specific rearrangement involved, 10 µl of the DNA deriving from eluted bands were amplified.

4.8. *Patient specific-probe design and MRD evaluation*

Allele-Specific Oligonucleotides (ASO) were designed by Primer Express® Software v2.0 (Applied Biosystems, Foster City, CA, USA) based on V-D-J junction regions unique for each lymphocyte rearrangement. For each rearrangement, at least two ASO forward primers were designed (version A or B) in the inserted or deleted junctional patient-specific regions. Therefore, ASO forward primers were used in combination with reverse primers and FAM-TAMRA probes targeting germline regions of the corresponding rearrangement, to be employed for MRD assessment by quantitative PCR (ASO-qPCR), since the sequence-specific amplification of the MRD-PCR products depends on the presence of leukemic cells.

According to the EuroMRD guidelines,[22,23] the probe sensitivity and specificity test was performed using a patient-specific standard curve made by 10-fold dilutions starting from the patient's 0.1 µg/µl DNA sample at diagnosis. Dilutions were prepared using MNCs-derived DNA, obtained from a pooled buffy-coat from 6-8 healthy blood donors. For the ASO test, two replicates of each dilution (from 10^{-2} to 10^{-5}), MNCs as No Amplification Control (NAC) and No Template Control (NTC) were tested. For each sample 500 ng of DNA per reaction were tested in a total volume of 25 µl in a mix solution as follows:

Reagent	**Initial concentration**	**Final concentration**	**Final volume**
TaqMan® Universal Master Mix	2X	1X	12.5 µl
ASO Forward primer	5 µM	300 nM	1.5 µl
Reverse primer	5 µM	300 nM	1.5 µl
Probe	5 µM	200 nM	1 µl
Bovine Serum Albumin (BSA)	20%	0.04%	0.05 µl
Nuclease free-water	-	-	3.45 µl
DNA	100 ng/µl	20 ng/µl	5 µl

Samples and mix were distributed in MicroAmp™ Optical 96-Well Reaction Plate with Barcode or MicroAmp™ Fast Optical 96-Well Reaction Plate with Barcode, 0.1 mL (Applied Biosystems, Foster City, CA, USA) and ASO-qPCR were performed using 7900HT Fast Real-Time PCR System instrument or ViiA™ 7 Dx Real-Time PCR System with 96-Well Block instrument, respectively, according to the following thermal protocol:

Temperature	Time	Cycles
50°C	2’	1
95°C	10’	1
95°C	15”	50
60°C*	60”	1
4°C	∞	1

The quality-check of each ASO-qPCR assay was performed, evaluating the assay sensitivity (indicating the lowest detectable MRD level, although not reproducibly and accurately) and the "quantitative range" (QR), reflecting the part of the standard curve in which the MRD level can be quantified reproducibly and accurately. ASO-qPCR assays reaching a QR of at least 10^{-4} passed the quality-check, while for less sensitive assays the real-time PCR thermal conditions were modified increasing the annealing temperature (*) of 1.5°C from 60°C to a maximum temperature of 69°C, in order to apply more stringent amplification conditions and enhance sensitivity and specificity of the test. For each patient, one or possibly two ASO primers were selected based on the assay performance for follow-up MRD measurement.

The ASO-qPCR for MRD evaluation was performed for each patient in triplicate for the target gene and in duplicate for the reference gene (albumin), using 500 ng of DNA per reaction and a total amount of 2,5 µg of DNA for each assay. The reference gene was tested as previously described using a standard curve made of 10-fold dilutions from 10^{-1} to 10^{-3} of MNCs-derived DNA dissolved in TE 1X. The annealing temperature for the reference gene amplification was 60°C. On the contrary, the target gene was tested for DNA samples in triplicate and standard curve dilutions in duplicate. Moreover, normal mononuclear cells (NAC) were run at least in 6-fold in each qPCR assay to evaluate the presence of aspecific amplification of the assay on normal lymphocytes. The annealing temperature for each target gene analysis was established based on the ASO-qPCR setting-up. The MRD results were expressed as the logarithmic reduction of the level of leukemic cells compared to the diagnosis.

5. RESULTS

5.1. *Validation of capture-based NGS panel: the panel identified all the Ig/TCR rearrangements found with the conventional method.*

The Ig/TCR identification in acute lymphoblastic leukemias is instrumental for Minimal Residual Disease (MRD) monitoring during disease treatment. Despite this process has been developed and ruled during the last 20 years within a European collaborative network (EuroClonality and EuroMRD within the ESHLO foundation) it is laborious, time-consuming and, in some cases, ineffective. With the aim to develop a simplified and possibly more efficient method, we designed a capture-based NGS approach. However, before applying this approach to prospective patients, this technology has to be proved at least equally effective as standard procedure. For this purpose, we started a validation process, selecting 10 diagnostic samples in which the conventional method identified overall 51 rearrangements. The capture-based NGS clonality assessment applied to these samples allowed to identify all the 51 known Ig/TCR clones shown in **Table 11**.

Interestingly, this NGS approach allowed to identify 24 additional clonal rearrangements involving the following loci: IGH (n=4), IGκ (n=2), IGλ (n=1), TCRβ (n=4), TCRγ (n=2), TCRδ and TCRα (n=11) **(Table 12)**. The majority of newly identified clonalities (n=11) was characterized by uncommon V-D-J combinations, not included into the standard PCR set designed to amplify only the most frequent rearrangements.[28,60] To exclude the possibility that the uncommon rearrangements derived from technical artifacts, we confirmed these rearrangements by qualitative PCR set up using specific oligonucleotides designed based on NGS-derived sequences, followed by Sanger sequencing (n=5 in TCRα, n=1 in TCRδ and n=5 in αδ loci). All the uncommon rearrangements were effectively amplified in the diagnostic samples and the sequences obtained by Sanger method on isolated standard amplification bands corresponded to the NGS sequences. Among the 24 newly identified clonalities, 10 were oligoclonal rearrangements that could not be solved by heteroduplex clonality assessment and Sanger sequencing. Indeed, we finally discriminated the rearrangements' specific sequences in Sanger electropherograms only thanks to single NGS-based sequences (single reads) (n=2 in IGκ, n=3 in IGH, n=3 in TCRβ and n=2 in TCRγ loci) **(Figure 12)**. Furthermore, 3 clonal rearrangements (n=1 IGH, n=1 TCRβ, and n=1 IGλ) were low represented clones that were missed by low sensitive standard PCR and finally confirmed in the diagnostic samples by re-amplification of faint heteroduplex amplicons. Without the *a priori* knowledge of these minor clones' presence, it has been impossible to recognize them at the time of screening.

As expected, no Ig/TCR rearrangements were revealed by NGS data analysis in HUVECs and mesenchymal cord blood cells we included as negative controls, proving that no procedure artifacts were introduced during libraries preparation, sequences generation or data analysis.

Patient ID	ALL lineage	Locus	V gene	del V	N	del D	D gene	del D	N	del J	J gene
BG_371	B	IGH	IGHV3-9*01	-2	15	-2	IGHD6-19*01	-4	7	0	IGHJ3*02
		TRD+	TRDV2*02	-4	10	-28	TRDD3*01	NA	NA	NA	NA
		TRG	TRGV2*01	0	20	NA	NA	NA	NA	-3	TRGJ1*01
		TRG	TRGV3*02	-4	15	NA	NA	NA	NA	-6	TRGJ1*02
BG_4502	T	TRB	TRBV11-2*03	0	9	-3	TRBD2*01	-5	1	-4	TRBJ2-1*01
		TRB	TRBV14*02	-3	4	-1	TRBD2*01	NA	NA	NA	TRBJ2-6*01
		TRD	TRDV1*01	0	15	0	TRDD3*01	-2	1	-3	TRDJ1*01
		TRG	TRGV3*02	-1	10	NA	NA	NA	NA	-4	TRGJ1*02
		TRG	TRGV9*01	-1	13	NA	NA	NA	NA	-6	TRGJ1*02
BG_5038	B	IGH	IGHV3-49*02	-15	8	-2	IGHD2-8*01	0	2	-2	IGHJ6*02
		IGK	IGKV1-16*02	-17	9	-6	KDE	NA	NA	NA	NA
		TRB	TRBV20-1*05	0	7	-3	TRBD2*01	-5	3	-2	TRBJ2-3*01
		TRD+	TRDD2*01	-6	5	NA	NA	NA	NA	0	TRDD3*01
		TRG	TRGV9*01	0	5	NA	NA	NA	NA	-1	TRGJP1*01
		TRG	TRGV3*02	0	4	NA	NA	NA	NA	0	TRGJ1*02
BG_5418	B	IGH	IGHV4-30-2*01	-14	34	0	IGHD3-3*01	-7	8	-11	IGHJ6*02
		IGH	IGHV3-7*02	-1	2	-10	IGHD2-2*01	-5	6	-4	IGHJ6*02
		IGK	IGKV3-20*01	-3	3	NA	NA	NA	NA	0	KDE
		IGL	IGLV2-8*01	-9	8	NA	NA	NA	NA	0	IGLJ2*01
		TRG	TRGV9*01	-10	10	NA	NA	NA	NA	-5	TRGJ1*02
		TRG	TRGV11*01	-7	8	NA	NA	NA	NA	-3	TRGJ1*02

BG_5452	T	IGH	IGHV3-11*01	0	19	-2	IGHD6-19*01	-4	4	-9	IGHJ4*02
		IGH	IGHV3-64D*06	-2	33	2	IGHD6-19*01	-4	4	9	IGHJ4*02
		IGH	IGHV4-34*01	0	6	1	IGHD6-6*01	-7	1	0	IGHJ6*02
		IGH+	NA	NA	NA	NA	IGHD7-27*01	-2	3	-3	IGHJ2*01
		TRB	NA	NA	NA	NA	TRBD1	33	0	0	TRBJ2-3
		TRB	TRBV9*01	-3	10	NA	TRBD2*01	NA	NA	0	TRBJ2-1*01
		TRG	TRGV10*02	-2	11	NA	NA	NA	NA	-6	TRGJ1*02
		TRG	TRGV2*01	0	3	NA	NA	NA	NA	-1	TRGJ1*02
BG_9574	T	TRB	TRBV14*01	-1	11	-4	NA	NA	NA	NA	TRBJ2-7*01
		TRB	TRBV5-3*01	-2	3	0	TRBD1*01	-3	27	-4	TRBJ2-7*01
		TRD	TRDV1*01	-5	5	0	TRDD2*01	0	11	0	TRDJ1*01
		TRG	TRGV11*01	-11	3	-2	NA	NA	NA	NA	TRGJ2*01
		TRG	TRGV4*02	-6	2	-1	NA	NA	NA	NA	TRGJ1*02
		TRG	TRGV8*01	-1	3	-3	NA	NA	NA	NA	TRGJP2*01
BG_9813	T	TRB+	NA	NA	NA	NA	TRBD2*01	-3	0	-5	TRBJ2-5*01
BG_9813	T	TRD+	TRDD2*01	-10	22	NA	NA	NA	NA	0	TRDJ1*01
		TRG	TRGV10*02	-1	3	NA	NA	NA	NA	-6	TRGJ1*02
		TRG	TRGV9*01	-3	3	NA	NA	NA	NA	0	TRGJ2*01
BG_11360	B	IGH	IGHV4-4*07	-2	9	-1	IGHD2-15*01	-6	12	-30	IGHJ6*04
		IGH+	NA	NA	NA	NA	IGHD7-27*01	0	12	-7	IGHJ6*02
		TRD+	TRDD2*01	0	5	NA	NA	NA	NA	-1	TRDD3*01
		TRD+	TRDV2*02	0	13	NA	NA	NA	NA	-3	TRDD3*01

BG_11720	B	IGH	IGHV6-1*01	0	5	-3	IGHD3-3*01	-2	2	-6	IGHJ6*02
		TRA+D	TRDV2*03	0	5	-2	TRDD3*01	0	3	-4	TRAJ9*01
		TRG	TRGV9*01	-6	13	NA	NA	NA	NA	-1	TRGJ1*02
BG_11806	B	IGH	IGHV1-2*02	-1	3	-3	IGHD2-21*02	-3	24	-1	IGHJ5*02
		IGK	IGKV2-28*01	0	6	NA	NA	NA	NA	-1	IGKJ4*02
		IGL	IGLV3-19*01	-5	14	NA	NA	NA	NA	-4	IGLJ3*02
		TRD+	TRDD2*01	0	2	NA	NA	NA	NA	0	TRDD3*01
		TRG	TRGV5*01	0	3	NA	NA	NA	NA	-1	TRGJ1*02

Table 11. *Ig/TCR rearrangements identified by NGS in the cohort of adult ALL patients enrolled into the NILG-ALL 09/00 clinical trial, formerly evaluated for clonality assessment. These rearrangements have been previously detailed by standard clonality assessment and Sanger sequencing following the conventional EuroMRD guidelines. IGK indicates IGκ rearrangements, IGL indicates IGλ rearrangements, TRB indicates TCRβ rearrangements, TRG indicates TCRγ rearrangements, TRD indicates TCRδ rearrangements and TRA indicates TCRα rearrangements. Symbol "+" refers to incomplete rearrangements.*

del V *Nucleotides deletion in V*
N *Number of nucleotides inserted*
del D *Nucleotides deletion in D*
del J *Nucleotides deletion in J*
NA *Not Applicable*
Y *Rearrangement already identified by standard procedure*

Patient ID	ALL lineage	Locus	V gene	del V	N	del D	D gene	del D	N	del J	J gene	Known	Rearrangement feature
BG_11806	B	TRD+	TRDD2*01	0	13	NA	NA	NA	NA	0	TRDJ3*01	N	Uncommon V/DJ combinations
BG_11806	B	TRA+D	TRDD2*01	-2	3	NA	NA	NA	NA	-7	TRAJ30*01	N	Uncommon V/DJ combinations
BG_5038	B	TRA+D	TRDD2*01	-3	53	NA	NA	NA	NA	-15	TRAJ48*01	N	Uncommon V/DJ combinations
BG_5418	B	TRA+D	TRDD2*01	-1	14	NA	NA	NA	NA	-5	TRAJ29*01	N	Uncommon V/DJ combinations
BG_5418	B	TRA+D	TRDD2*01	-11	4	NA	NA	NA	NA	-1	TRAJ23*01	N	Uncommon V/DJ combinations
BG_5418	B	TRA	TRAV26-1*01	0	4	NA	NA	NA	NA	-2	TRAJ33*01	N	Uncommon V/DJ combinations
BG_5418	B	TRA	TRAV8-3*01	0	8	NA	NA	NA	NA	-7	TRAJ34*01	N	Uncommon V/DJ combinations
BG_5452	T	TRA	TRAV21*01	0	1	NA	NA	NA	NA	0	TRAJ48*01	N	Uncommon V/DJ combinations
BG_9574	T	TRA+D	TRAV29/DV5*01	0	38	0	TRDD3*01	-4	3	0	TRDJ1*01	N	Uncommon V/DJ combinations
BG_9574	T	TRA	TRAV21*02	-7	0	-5	NA	NA	NA	NA	TRAJ24*02	N	Uncommon V/DJ combinations
BG_9813	T	TRA	TRAV19*01	-2	10	NA	NA	NA	NA	-4	TRAJ36*01	N	Uncommon V/DJ combinations
BG_11360	B	IGH	IGHV4-34*01	0	9	-18	IGHD2-2*01	0	2	-4	IGHJ6*03	N	Oligoclonality
BG_11720	B	IGH	IGHV4-34*01	0	6	-1	IGHD6-6*01	-7	1	0	IGHJ6*02	N	Oligoclonality
BG_371	B	IGK	IGKV1-33*01	-1	2	NA	NA	NA	NA	-2	IGKJ4*01	N	Oligoclonality
BG_371	B	IGK	IGKV1-39*01	-4	8	NA	NA	NA	NA	-9	IGKJ2*02	N	Oligoclonality
BG_371	B	TRG	TRGV4*02	-4	5	NA	NA	NA	NA	-3	TRGJ1*01	N	Oligoclonality
BG_5418	B	TRB	TRBV23-1*01	0	1	0	TRBD2*01	-7	14	-6	TRBJ2-7*01	N	Oligoclonality
BG_5418	B	TRB	TRBV10-3*01	-8	18	-7	TRBD2*01	NA	NA	NA	TRBJ2-3*01	N	Oligoclonality
BG_5418	B	TRB	TRBV24-1*01	-13	5	-2	TRBD1*01	-3	5	2	TRBJ2-7*01	N	Oligoclonality
BG_5452	T	IGH	IGHV6-1*01	0	5	-3	IGHD3-3*01	-2	2	-6	IGHJ6*02	N	Oligoclonality
BG_9574	T	TRG	TRGV11*01	0	1	-25	NA	NA	NA	NA	TRGJ1*02	N	Oligoclonality

BG_11360	B	IGL	IGLV3-10*01	-2	7	NA	NA	NA	NA	0	IGLJ3*02	N	Low represented clone
BG_371	B	IGH	IGHV1-3*02	-2	4	-17	IGHD3-16*01	-15	0	-18	IGHJ4*02	N	Low represented clone
BG_9813	T	TRB	TRBV4-3*01	-1	21	-6	TRBD2*02	NA	NA	NA	TRBJ1-1*01	N	Low represented clone

Table 12. *Additional Ig/TCR rearrangements identified by NGS in the cohort of adult ALL patients enrolled into the NILG -ALL 09/00 clinical trial, formerly evaluated for clonality assessment. These rearrangements have not been previously identified by standard clonality assessment and Sanger sequencing following the conventional EuroMRD guidelines. IGK indicates IGκ rearrangements, IGL indicates IGλ rearrangements, TRB indicates TCRβ rearrangements, TRG indicates TCRγ rearrangements, TRD indicates TCRδ rearrangements and TRA indicates TCRα rearrangements. Symbol "+" refers to incomplete rearrangements.*

del V *Nucleotides deletion in V*
N *Number of nucleotides inserted*
del D *Nucleotides deletion in D*
del J *Nucleotides deletion in J*
NA *Not Applicable*
N *Rearrangement identified by NGS, not previously identified by standard procedure*

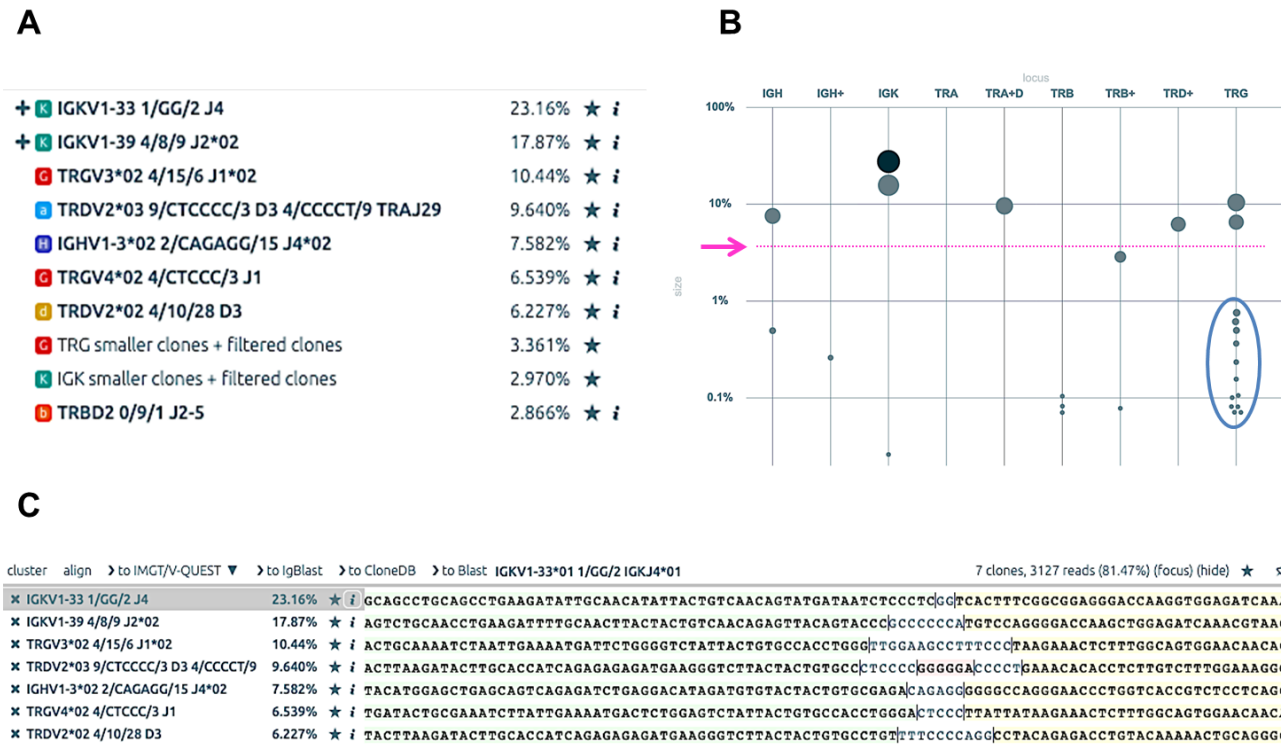

Figure 12. *Visualization of NGS-derived data analyzed by Vidjil software. The Vidjil tool is based on an algorithm gathering NGS-derived reads into clonotypes, allowing to visualize and analyze clones and tracking them along the time in an MRD setup or immunological studies. The algorithm works on reads coming from either amplicon-based or capture-based deep sequencing strategy, as soon as they include CDR3 sequences. The web application displays Ig/TCR clonal rearrangements (clonotypes) on a grid (A) with their relative clone abundance in the sample, also reported in a bubble chart (B), and on a list with their representative sequences, possibly aligned to the reference human genome (C). Clonotypes can be manually annotated, edited or clustered. Moreover, it is possible to plot on the bubble chart several metrics on the lymphocyte population. Thanks to the NGS approach we were able to discriminate oligoclonal rearrangements (panel B, blue circle), belonging to the same gene family, that could not be solved by heteroduplex clonality assessment and Sanger sequencing. Furthermore, the NGS approach allowed the recognition of minor leukemia clones, below the sensitivity limit of standard PCR (panel B, red line).*

5.2. *MRD quantification by capture-based NGS panel*

Standard MRD evaluation on follow-up samples is performed by Real time-qPCR using a rearrangement-specific oligonucleotide in combination with family-specific primers and probes. The quantification of residual disease is obtained comparing the amplification plot of the follow-up samples to a standard curve consisting of 10-fold dilutions of diagnostic material in normal cells, mimicking the leukemia reduction during treatment. The MRD level is expressed as the logarithmic reduction compared to diagnosis. Despite this method proved sensitive and useful to predict leukemia relapse, some low positivity cannot be distinguished from non-specific signals caused by miss-annealing of specific oligonucleotides to similar, normal sequences.

NGS approach could overcome this limitation because single sequences can be obtained and compared to the leukemic clone sequence. Therefore, for each of the 10 above mentioned patients, we also tested 3 follow-up samples by NGS. The follow-up MRD level by ASO-qPCR ranged from 10^{-2} to negative, including very low positivity (10^{-5} or positive not quantifiable) samples. Unfortunately, none of the 30 follow-up samples provided evidence of clonal rearrangements at any level. To exclude that the negativity was due to a low DNA input, we performed an NGS experiment on one additional diagnostic sample and its 10-fold dilution curve, using the maximum DNA input (1 μg) allowed by the assay.

In the undiluted diagnostic sample, we identified 6 rearrangements with a relative percentage ranging from 21.6 to 6.6%. All these rearrangements were identified also in 10^{-1} and 10^{-2} dilution with the expected relative clone reduction, but only two rearrangements were still detectable in 10^{-3} dilution. We subsequently also analyzed by NGS approach two follow-up samples of the same patient. These samples had been previously studied by ASO-qPCR assays designed on two different clones and had been concordantly defined as MRD positive at 10^{-2} level and positive, not quantifiable ($<10^{-4}$). In the first follow-up sample, we recognized by NGS all the rearrangements identified at diagnosis with a percentage ranging from 5 to 1%. In particular, the two rearrangements selected for patient-specific molecular probes were identified at 1.3% and 1%, respectively, in accordance with the 10^{-2} positivity. In the second follow-up sample, the NGS assay did not recognize any specific rearrangement.

5.3. *The retrospective application of capture-based NGS panel to ALL patients lacking a recognized clonality at diagnosis (MRD-unknown) by standard method allowed to identify useful Ig/TCR rearrangements*

In the group of 23 MRD-unknown ALL patients, for which it had been impossible to isolate clonal rearrangements or to obtain a sensitive patient-specific molecular probe at the time of diagnosis, we detected at least one rearrangement in 20 out of 23 patients (87%). In all the 13 adult B-ALL cases, we recognized at least one V-D-J or D-J clonal rearrangement. On the contrary, in adult T-ALL cases our approach identified V-D-J or D-J clonal rearrangements in 7 out of 10 cases, while it could not reveal any clonal rearrangement in the remaining 3 samples.

Overall, the capture-based NGS allowed to identify 92 clonal rearrangements (72 in B- ALL, 20 in T-ALL) in IGH (n=30), IGκ (n=6), TCRβ (n=6), TCRγ (n=32), TCRδ (n=9), TCRαδ (n=3) and TCRα (n=6) loci, detailed in **Table 13**.

Then, we proceeded to the validation of these results with conventional methods. Sixty-nine rearrangements defined by NGS with a level ≥5% underwent the validation process by the conventional method. Among these cases, 53 out of 69 were amplified and Sanger sequenced successfully using implemented standard PCR methods, based on the updated EuroMRD guidelines, or specific PCR appropriately designed on NGS sequences. In the remaining 16 rearrangements, the validation failed since NGS sequences were not confirmed by conventional Sanger which did not allow to discriminate oligoclonality nor cases with a low clone representation (slightly higher than the sensitivity limit of 5%). Twenty-one rearrangements were found by NGS at a level below 5%, so representing minor leukemia clones. This level is out of the sensitivity of the standard approach, so preventing the possibility to proceed to a validation process. Two additional clonal rearrangements identified in one sample, could not be validated because of the exhaustion of the diagnostic DNA.

Patient ID	ALL Lineage	Locus	V gene	del V	N	del D	D gene	del D	N	del J	J gene	Validation	Rearrangement feature
BG_11584	B	IGH	IGHV6-1*01	-2	8	-6	IGHD6-6*01	0	3	-10	IGHJ4*02	Y	
BG_2097	B	IGH	NA	NA	NA	NA	IGHD4-23*01	-2	1	-4	IGHJ2*01	Y	
BG_855	B	IGH	IGHV6-1*01	-3	2	-4	IGHD2-2*01	-1	5	0	IGHJ6*03	Y	
BG_855	B	IGH	IGHV3-23*01	0	19	-2	IGHD3-9*01	-8	13	-7	IGHJ4*02	Y	
BG_10112	B	IGH+	NA	NA	NA	NA	IGHD6-6*01	-2	0	-7	IGHJ4*02	Y	
BG_10112	B	IGH+	NA	NA	NA	NA	IGHD1-7*01	-6	2	-4	IGHJ4*02	Y	
BG_11053	B	IGH+	NA	NA	NA	NA	IGHD2-2*01	-6	3	-4	IGHJ4*02	Y	
BG_1125	B	IGH+	NA	NA	NA	NA	IGHD6-25*01	0	-5	11	IGHJ4*02	Y	
BG_11584	B	IGH+	NA	NA	NA	NA	IGHD1-26*01	0	3	-3	IGHJ3*02	Y	
BG_4254	B	IGH+	NA	NA	NA	NA	IGHD3-9*01	-5	4	-11	IGHJ6*03	Y	
BG_4254	B	IGH+	NA	NA	NA	NA	IGHD6-6*01	-4	3	-15	IGHJ5*02	Y	
BG_5702	B	IGH+	NA	NA	NA	NA	IGHD2-2*02	-6	7	-5	IGHJ5*02	Y	
BG_8345	B	IGH+	NA	NA	NA	NA	IGHD3-22*01	-2	7	1	IGHJ6*02	Y	
BG_8345	B	IGH+	NA	NA	NA	NA	IGHD1-26*01	-2	0	7	IGHJ4*02	Y	
BG_9445	B	IGH+	NA	NA	NA	NA	IGHD2-2*02	-3	19	5	IGHJ6*03	Y	
BG_11269	T	TRB	TRBV7-8*01	-3	14	NA	NA	NA	NA	-7	TRBJ1-4*01	Y	
BG_6037	T	TRB	TRBV4-2*01	0	25	NA	NA	NA	NA	-3	TRBJ2-3*01	Y	
BG_6037	T	TRB	TRBV20-1*02	-3	15	-2	TRBD2*02	-2	19	-6	TRBJ2-1*01	Y	
BG_10112	B	TRG	TRGV11*01	-4	7	NA	NA	NA	NA	-8	TRGJP1*01	Y	
BG_10112	B	TRG	TRGV11*01	-3	10	N A	NA	NA	NA	-1	TRGJ2*01	Y	
BG_11269	T	TRG	TRGV4*02	-4	5	NA	NA	NA	NA	0	TRGJ2*01	Y	

BG_11269	T	TRG	TRGV10*02	-6	0	NA	NA	NA	NA	-8	TRGJ1*02	Y	
BG_11584	B	TRG	TRGV4*02	-6	2	NA	NA	NA	NA	0	TRGJ2*01	Y	
BG_11584	B	TRG	TRGV10*02	0	NA	NA	NA	NA	4	-9	TRGJP1*01	Y	
BG_11584	B	TRG	TRGV11*01	0	0	NA	NA	NA	NA	-7	TRGJP1*01	Y	
BG_855	B	TRG	TRGV9*01	-10	3	NA	NA	NA	NA	-2	TRGJ1*02	Y	
BG_11053	B	TRD+	NA	NA	NA	NA	TRDD2*01	-1	22	-1	TRDJ1*01	Y	
BG_11053	B	TRD+	NA	NA	NA	NA	TRDD2*01	0	19	0	TRDJ4*01	Y	
BG_4254	B	TRD+	NA	NA	NA	NA	TRDD2*01	0	5	-2	TRDD3*01	Y	
BG_8345	B	TRD+	NA	NA	NA	NA	TRDD2*01	-3	15	0	TRDD3*01	Y	
BG_11053	B	IGH+	NA	NA	NA	NA	IGHD3-9*01	-5	3	-4	IGHJ4*02	Y	Oligoclonality
BG_11053	B	IGH+	NA	NA	NA	NA	IGHD3-9*01	-5	0	-4	IGHJ6*02	Y	Oligoclonality
BG_11053	B	IGH+	NA	NA	NA	NA	IGHD3-9*01	-3	13	-6	IGHJ6*02	Y	Oligoclonality
BG_11345	T	TRD+	NA	NA	NA	NA	TRDD2*01	-8	0	-14	TRDD3*01	Y	Oligoclonality
BG_1125	B	IGK	IGKV1-27*01	-1	0	NA	NA	NA	NA	2	IGKJ1*01	Y[F]	
BG_11269	T	IGK	IGKV4-1*01	-1	0	NA	NA	NA	NA	-4	IGKJ4*01	Y[F]	
BG_11269	T	IGK	IGKV2D-29*01	-2	6	NA	NA	NA	NA	0	IGKJ2*01	Y[F]	
BG_11269	T	IGK	IGKV1-6*01	-3	0	NA	NA	NA	NA	0	IGKJ1*01	Y[F]	
BG_6037	T	TRG	TRGV3*01	0	0	NA	NA	NA	NA	0	TRGJ1*02	Y[F]	
BG_6037	T	TRG	TRGV2*01	0	4	NA	NA	NA	NA	-2	TRGJ1*02	Y[F]	
BG_2097	B	TRA+D	TRDV1*01	-2	8	NA	NA	NA	NA	-8	TRAJ29*01	Y[F]	
BG_855	B	TRA+D	NA	NA	NA	NA	TRDD2*01	-4	24	-5	TRAJ29*01	Y[F]	
BG_855	B	TRA+D	TRDV2*01	0	6	0	TRDD3*01	0	5	-4	TRAJ58*01	Y[F]	

BG_11269	T	TRA	TRAV26-1*01	-3	3	NA	NA	NA	NA	-4	TRAJ4*01	Yᵀ	
BG_12438	T	TRA	TRAV19*01	-8	4	NA	NA	NA	NA	0	TRAJ47*02	Yᵀ	
BG_2097	B	TRA	TRAV13-1*01	-3	0	NA	NA	NA	NA	-3	TRAJ35*01	Yᵀ	
BG_5702	B	TRA	TRAV16*01	-6	2	NA	NA	NA	NA	-4	TRAJ9*01	Yᵀ	
BG_10487	B	IGH+	NA	NA	NA	NA	IGHD3-16*02	-5	8	0	IGHJ4*02	Yᵀ	Biclonal sequence
BG_10487	B	IGH+	NA	NA	NA	NA	IGHD3-3*01	0	11	-3	IGHJ5*02	Yᵀ	Biclonal sequence
BG_6490	B	IGH+	NA	NA	NA	NA	IGHD2-21*02	-1	2	0	IGHJ6*03	Yᵀ	Biclonal sequence
BG_6490	B	IGH+	NA	NA	NA	NA	IGHD2-8*01	-7	8	-2	IGHJ3*02	Yᵀ	Biclonal sequence
BG_8646	T	TRD+	NA	NA	NA	NA	TRDD2*01	-8	0	-14	TRDD3*01	Yᵀ	Biclonal sequence
BG_8646	T	TRD+	NA	NA	NA	NA	TRDD2*01	-8	0	-16	TRDD3*01	Yᵀ	Biclonal sequence
BG_1125	B	IGH	IGHV4-34*02	0	10	NA	NA	NA	NA	-4	IGHJ5*01	N	Low represented clone
BG_5702	B	IGK	IGKV2-28*01	-3	2	-7	KDE	NA	NA	NA	NA	N	Low represented clone
BG_12438	T	TRB	TRBV4-1*02	-3	2	-1	TRBD1*01	-3	8	-2	TRBJ2-7*01	N	Low represented clone
BG_5702	B	TRB	TRBV4-1*02	0	2	0	TRBD1*01	-4	10	-8	TRBJ1-1*01	N	Low represented clone
BG_11053	B	TRG	TRGV9*01	-1	6	NA	NA	NA	NA	0	TRGJ1*02	N	Low represented clone
BG_3895	T	TRG	TRGV2*01	0	4	NA	NA	NA	NA	-10	TRGJP2*01	N	Low represented clone
BG_11053	B	TRD+	NA	NA	NA	NA	TRDD2*01	0	14	-10	TRDJ1*01	N	Low represented clone
BG_12438	T	TRA	TRAV21*01	-4	7	NA	NA	NA	NA	0	TRAJ27*01	N	Low represented clone
BG_2097	B	TRA	TRAV19*01	-10	5	NA	NA	NA	NA	0	TRAJ47*01	N	Low represented clone
BG_6490	B	IGH	IGHV4-34*12	0	9	1	IGHD2-8*02	-7	8	-2	IGHJ3*02	N	Oligoclonality
BG_5702	B	IGK	IGKV3-20*01	-4	6	NA	NA	NA	NA	-2	IGKJ2*01	N	Oligoclonality
BG_2097	B	TRB	TRBV6-5*01	0	2	-4	TRBD2*01	-7	0	-4	TRBJ1-5*01	N	Oligoclonality

BG_11053	B	TRG	TRGV4*02	-3	0	NA	NA	NA	NA	-4	TRGJ2*01	N	Oligoclonality
BG_11053	B	TRG	TRGV11*01	-7	7	NA	NA	NA	NA	-1	TRGJ2*01	N	Oligoclonality
BG_2097	B	TRG	TRGV3*01	0	15	NA	NA	NA	NA	-1	TRGJ2*01	N	Oligoclonality
BG_9445	B	TRG	TRGV10*02	0	3	NA	NA	NA	NA	4	TRGJ1*01	N	Oligoclonality
BG_11584	B	IGH	IGHV4-31*02	-1	1	-3	IGHD3-10*01	-6	0	0	IGHJ4*02	ND	<5%
BG_2481	T	IGH	NA	NA	NA	NA	IGHD7-27*01	-11	0	0	IGHJ1*01	ND	<5%
BG_11054	B	IGH+	NA	NA	NA	NA	IGHD6-13*01	3	0	21	IGHJ1*01	ND	<5%
BG_11054	B	IGH+	NA	NA	NA	NA	IGHD7-27*01	0	2	6	IGHJ4*02	ND	<5%
BG_10112	B	TRG	TRGV5*01	0	6	NA	NA	NA	NA	-1	TRGJ1*02	ND	<5%
BG_10112	B	TRG	TRGV9*01	0	9	NA	NA	NA	NA	0	TRGJ1*02	ND	<5%
BG_10112	B	TRG	TRGV11*01	-9	12	NA	NA	NA	NA	-5	TRGJ1*02	ND	<5%
BG_10112	B	TRG	TRGV9*01	-2	4	NA	NA	NA	NA	0	TRGJ2*01	ND	<5%
BG_10112	B	TRG	TRGV9*01	0	4	NA	NA	NA	NA	-2	TRGJ2*01	ND	<5%
BG_11053	B	TRG	TRGV10*02	-2	0	NA	NA	NA	NA	-1	TRGJ2*01	ND	<5%
BG_11053	B	TRG	TRGV11*01	-2	15	NA	NA	NA	NA	0	TRGJ1*02	ND	<5%
BG_11053	B	TRG	TRGV11*01	0	10	NA	NA	NA	NA	-8	TRGJ1*01	ND	<5%
BG_11053	B	TRG	TRGV5*01	0	3	NA	NA	NA	NA	-5	TRGJ1*02	ND	<5%
BG_11584	B	TRG	TRGV11*01	0	NA	NA	NA	NA	5	-2	TRGJ1*02	ND	<5%
BG_11584	B	TRG	TRGV11*01	-7	NA	NA	NA	NA	3	-7	TRGJP1*01	ND	<5%
BG_11584	B	TRG	TRGV9*01	-2	NA	NA	NA	NA	11	-8	TRGJ1*01	ND	<5%
BG_11584	B	TRG	TRGV11*01	-3	NA	NA	NA	NA	0	-16	TRGJ1*02	ND	<5%
BG_2097	B	TRG	TRGV4*02	0	6	NA	NA	NA	NA	-5	TRGJP1*01	ND	<5%

BG_2097	B	TRG	TRGV4*01	-5	14	NA	NA	NA	NA	0	TRGJ1*02	ND	<5%
BG_11053	B	TRD+	NA	NA	NA	NA	TRDD2*01	0	16	0	TRDJ1*01	ND	<5%
BG_12438	T	TRG	TRGV8*01	1	0	NA	NA	NA	NA	-7	TRGJP2*01	ND	ND
BG_10640	B	IGH+	NA	NA	NA	NA	IGHD3-3*01	-4	6	-5	IGHJ6*02	ND•	
BG_10640	B	IGH+	NA	NA	NA	NA	IGHD2-2*02	-3	4	-10	IGHJ6*02	ND•	
BG_4005	T	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na		
BG_4255	T	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na		
BG_4379	T	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na		

Table 13. *Ig/TCR rearrangements identified by NGS in the MRD unknown cohort of adult ALL patients enrolled into the NILG -ALL 09/00 clinical trial. These rearrangements have not been previously identified by standard clonality assessment and Sanger sequencing following the conventional EuroMRD guidelines, or they were not adequate to generate a suitable patient-specific molecular probe for the MRD evaluation. IGK indicates IGκ rearrangements, IGL indicates IGλ rearrangements, TRB indicates TCRβ rearrangements, TRG indicates TCRγ rearrangements, TRD indicates TCRδ rearrangements and TRA indicates TCRα rearrangements. Symbol "+" refers to incomplete rearrangements. For patients BG_4005, BG_4255 and BG_4379, no Ig/TCR rearrangements were detected.*

del V *Nucleotides deletion in V*
N *Number of nucleotides inserted*
del D *Nucleotides deletion in D*
del J *Nucleotide deletion in J*
NA *Not applicable*
Na *Not available*
Y *Rearrangement confirmed by standard procedure*
T *Use of specific primers for conventional PCR set up required*
N *Rearrangement identified by NGS not validated by standard procedure*
ND *Not Done*
• *Exhausted diagnostic material*

5.4. *Incomplete IGH rearrangements prevail in MRD-unknown patients carrying TP53 mutations.*

In a previous study, the cohort of 23 MRD-unknown ALL patients underwent the analysis of *TP53* mutational status by an independent NGS approach.[61] Indeed, the impact of *TP53* alterations is not completely clarified in the context of ALL. The study performed in our laboratory demonstrated the occurrence of *TP53* mutations in 14 out of 171 (8%) patients enrolled into the NILG-ALL 09/00 clinical trial. Furthermore, the presence of pathogenic variants in *the TP53* gene proved related to an increased risk of relapse, while not responsible for primary resistance to induction chemotherapy thus preventing the achievement of Complete Remission (CR). Of note, it was not possible to correlate *TP53* mutational status and MRD, the most important prognostic factor in ALL, because in the majority of mutated patients, 9 out of 14 (64%), no informative Ig/TCR rearrangements were identified for the design of a suitable patient-specific molecular probe (MRD-unknown patients). As a matter of fact, the study demonstrated a statistically significant correlation between positivity to *TP53* mutations and the unavailability of Ig/TCR-based patient-specific probes ($p<0.0001$).

In the group of 9 MRD-unknown patients carrying *TP53* mutations included in our analysis, we not only identified at least one Ig/TCR rearrangement by NGS, but we also highlighted the prevalence of incomplete IGH rearrangements. Indeed, among this group, all the B-ALL cases (8 out of 9) were characterized by the presence of at least one IGH incomplete clonal sequence **(Table 14)**. Since the assessment of incomplete IGH rearrangements was not performed at the time of patients' enrollment, it has been impossible to detect at diagnosis these informative clonal markers. However, the implementation of standard PCR methods available in the updated EuroMRD guidelines allowed the successful validation of the new identified incomplete rearrangements. The remaining *TP53* mutated patient (BG_8646), in which no incomplete IGH rearrangements were found, was actually a T-lineage ALL patient. In this case, the missing of informative Ig/TCR clonal markers relied on the presence of biclonal rearrangements involving the same TCRδ locus, identified only thanks to single NGS-derived sequences **(Figure 13)**.

Figure 13. *Biclonal TRDD2*01-DD3*01 rearrangements, discriminated by single NGS sequences. For each NGS-derived sequence, it was possible to recognize the different gene segments (green and yellow) of the given rearrangement. By contrast, double peaks in the Sanger electropherogram prevent the correct attribution of nucleotides to the specific clonal sequence.*

Patient ID	ALL lineage	Genetics *TP53*	*TP53* Mutation	VAF (%)	Ig/TCR probe	Locus	D gene	del D	N	del J	J gene	Procedure
BG_4205#	B	MUT+DEL	p.R273C	79	N	IGH+	NT	NT	NT	NT	NT	ND•
BG_2873#	B	MUT	p.C238Y	45	N	IGH+	IGHD6-6*01	-1	6	-8	IGHJ6*02	Standard
BG_6490*	B	MUT	p.V274A	90	N	IGH+	IGHD2-21*02	-1	2	0	IGHJ6*03	NGS
						IGH+	IGHD2-8*01	-7	8	-2	IGHJ3*02	NGS
BG_5702*	B	MUT	p.P151T	80	N	IGH+	IGHD2-2*02	-6	7	-5	IGHJ5*02	NGS
BG_2097*	B	MUT	p.R267W	7	N	IGH+	IGHD4-23*01	-2	1	-4	IGHJ2*01	NGS
BG_8345*	B	MUT	DUP 7577085,,7577093	63	N	IGH+	IGHD3-22*01	-2	7	1	IGHJ6*02	NGS
						IGH+	IGHD1-26*01	-2	0	7	IGHJ4*02	NGS
BG_11584*	B	MUT	p.R282G	4	N; t(4;11)	IGH+	IGHD1-26*01	0	3	-3	IGHJ3*02	NGS
BG_4254*	B	MUT	p.R175H	97	N	IGH+	IGHD3-9*01	-5	4	-11	IGHJ6*03	NGS
						IGH+	IGHD6-6*01	-4	3	-15	IGHJ5*02	NGS
BG_10112*	B	MUT	p.S166L	37	N; t(4;11)	IGH+	IGHD6-6*01	-2	0	-7	IGHJ4*02	NGS
						IGH+	IGHD1-7*01	-6	2	-4	IGHJ4*02	NGS
BG_9445*	B	MUT+DEL	p.R247W	90	N	IGH+	IGHD2-2*02	-3	19	5	IGHJ6*03	NGS
BG_11543	B	MUT+DEL	p.R273H	22	Y; DD2DD3	IGH+	IGHD2-21*02	-5	16	-17	IGHJ6*02	Standard
BG_8646*	T	MUT	p.R273C	5	N	IGH+	Na	Na	Na	Na	Na	NGS
BG_8142	T	dMUT	p.P222L / p.R158H	85/80	Y; VGI	IGH+	Na	Na	Na	Na	Na	Standard
BG_10442	T	dMUT	DUP 7578192,,7578195/ 7577597INS(GGGGTTTGACC)	36/21	Y; JB2.3	IGH+	Na	Na	Na	Na	Na	NGS; Standard

Table 14. *Patients enrolled into the NILG -ALL 09/00 clinical trial proved positive for mutations in the TP53 gene. For three patients (BG_11543, BG_10442, and BG_8142) a patient-specific probe based on Ig/TCR rearrangements was designed during the marker screening phase at disease onset. * Patients belonging to the cohort of 23 MRD-unknown cases, since no Ig/TCR molecular markers were available by standard approach. # Patients BG_4205 and BG_2873, while MRD-unknown, have not been evaluated by NGS for the retrospective identification of Ig/TCR clonal rearrangements. Among the group of TP53-mutated patients, all the B-ALL cases tested were characterized by the presence of at least one IGH incomplete (IGH+) clonal sequence, while*

none of the 3 T-ALL patients harbored this type of rearrangement." dMUT" refers to the double mutation of TP53, "MUT+DEL" refers to presence of one TP53 mutation and deletion of second TP53 allele.

VAF *Variant Allele Fraction*
Y *Identification of clonal Ig/TCR rearrangement by standard procedure for MRD evaluation*
N *Lacking of clonal Ig/TCR rearrangement by standard procedure for MRD evaluation*
del D *Nucleotides deletion in D*
N *Number of nucleotides inserted*
del J *Nucleotide deletion in J*
Na *Not available*
ND *Not Done*
• *Exhausted diagnostic material*
NT *Not Tested*

Based on this preliminary result, we completed by standard approach the analysis of incomplete IGH rearrangements in the group of *TP53* mutated ALL cases not yet studied excluding patient BG_4205, for which diagnostic material was exhausted. The concomitant presence of *TP53* gene alterations and incomplete IGH rearrangements was detected overall in 10 out of 13 cases (77%). In particular, the concomitant presence was detected in all the 10 B-ALL patients carrying *TP53* molecular alterations (100%), while none of the 3 T-ALL patients proved positive for IGH-D clonal sequences.

Therefore, we tested the available residual diagnostic material from a group of 25 B-ALL patients within the NILG-ALL 09/00 trial, proved negative for *TP53* mutations **(Table 15)**. We identified incomplete IGH rearrangements only in 9 out of 25 (36%) *TP53*-negative B-ALL patients. Despite the limits of a small number of subjects analyzed, we confirmed the higher prevalence of incomplete IGH rearrangements in *TP53* mutated B-ALL patients, compared to the *TP53* wild type counterpart **(Figure 14)**.

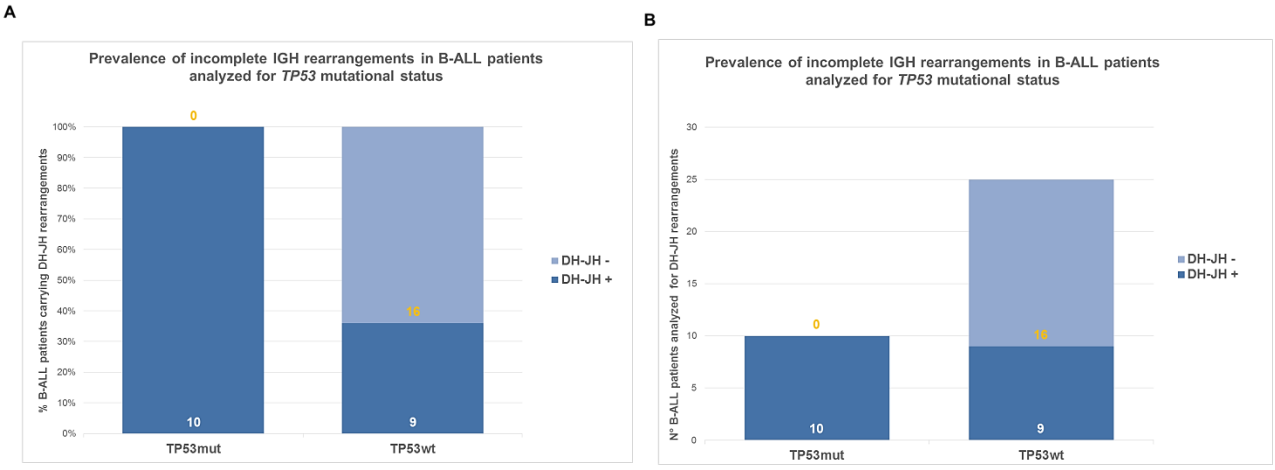

Figure 14. *Prevalence of incomplete (DH-JH) IGH rearrangements in B-ALL patients analyzed for the mutational status of the TP53 gene. Percentage (panel A) and absolute number (panel B) of B-ALL patients carrying incomplete IGH rearrangements in the group of TP53 mutated (mut) and TP53 wild type (wt) B-ALL patients. The absolute number of analyzed cases is indicated in white for the cohort of patients carrying incomplete IGH rearrangements (DH-JH +) and in yellow for the cohort of patients negative for incomplete IGH rearrangements (DH-JH -).*

Patient ID	ALL lineage	Genetics *TP53*	Ig/TCR probe	Locus	D gene	del D	N	del J	J gene	Procedure
BG_630	B	Wild Type	Y	IGH+	IGHD2-2*01	-3	9	-5	IGHJ6*02	Standard
BG_2195	B	Wild Type	Y	IGH+	IGHD3-9*01	-9	12	-2	IGHJ4*02	Standard
BG_5077	B	Wild Type	Y	IGH+	IGHD5-24*01	-2	1	-16	IGHJ5*02	Standard
BG_5136	B	Wild Type	Y	IGH+	IGHD2-15*01	-8	10	-14	IGHJ4*02	Standard
BG_7567	B	Wild Type	Y	IGH+	IGHD6-19*01	-6	21	-2	IGHJ6*02	Standard
BG_8632	B	Wild Type	Y	IGH+	IGHD4-17*01	0	7	-6	IGHJ5*02	Standard
BG_11360	B	Wild Type	Y	IGH+	IGHD7-27*01	0	-12	-7	IGHJ6*02	Standard
BG_11784	B	Wild Type	Y	IGH+	IGHD5-12*01	-2	3	-5	IGHJ6*02	Standard
BG_12374	B	Wild Type	Y	IGH+	IGHD2-2*01	-3	26	-4	IGHJ6*02	Standard
					IGHD3-3*01	-4	9	-4	IGHJ6*03	Standard
BG_1839	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_2145	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_2567	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_4612	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_4790	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_5956	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_6333	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_6380	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_6533	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_7275	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_8206	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard

BG_8844	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_10981	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_11568	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_11696	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard
BG_11958	B	Wild Type	Y	IGH+	Na	Na	Na	Na	Na	Standard

Table 15. *Patients affected by B-ALL enrolled into the NILG -ALL 09/00 clinical trial proved negative for mutations in the TP53 gene. For all the patients at least a patient-specific probe based on Ig/TCR rearrangements was designed during the marker screening phase at disease onset. Incomplete (+) IGH rearrangements were tested by standard procedure and detected in 9 out of 25 (36%) TP53-negative B-ALL cases.*

Y *Identification of clonal Ig/TCR rearrangement by standard procedure for MRD evaluation*
del D *Nucleotides deletion in D*
N *Number of nucleotides inserted*
del J *Nucleotide deletion in J*
Na *Not available*

5.5. *Retrospective application of capture-based NGS-derived sequences allowed the MRD evaluation in MRD-unknown ALL patients*

Based on NGS-derived sequences, we retrospectively designed ASO primers for 6 patients belonging to the MRD-unknown cohort, to further validate clonal rearrangements identification and to verify the correlation between NGS-clone percentage and clone abundance at diagnosis. In particular, we obtained 9 sensitive patient-specific molecular probes for 6 patients, reaching the required sensitivity of 10^{-4} or 10^{-5} and a quantitative range (QR) of 10^{-4} **(Figure 15)**. Therefore, we proceeded with the MRD assessment by ASO-qPCR in 4 of the above-mentioned patients with available follow-up samples. For each patient, we evaluated a minimum of 3 samples, including decisional time-point samples before cycle 4 (TP1), cycle 6 (TP2) and cycle 8 (TP3) of the treatment, according to the NILG-ALL 09/00 regimen, and eventually relapse samples (n=4) **(Figure 16)**. For all the studied patients we were able to properly follow the disease evolution, detecting very low leukemia levels (expressed as MRD positivity not quantifiable), in at least one decisional time-point **(Table 16)**. These results, if available at the time of the treatment, could have addressed the therapy to a more aggressive regimen and/or possibly HSCT, based on MRD positivity.

The MRD persistence data aligned to results obtained for the 3 patients (BG_11543, BG_10442, and BG_8142) likewise *TP53* mutated, for which a molecular patient-specific probe was available at the time of marker screening. For patient BG_10442 MRD evaluation for TP2 and TP3 was not possible because the patient relapsed soon after TP1.

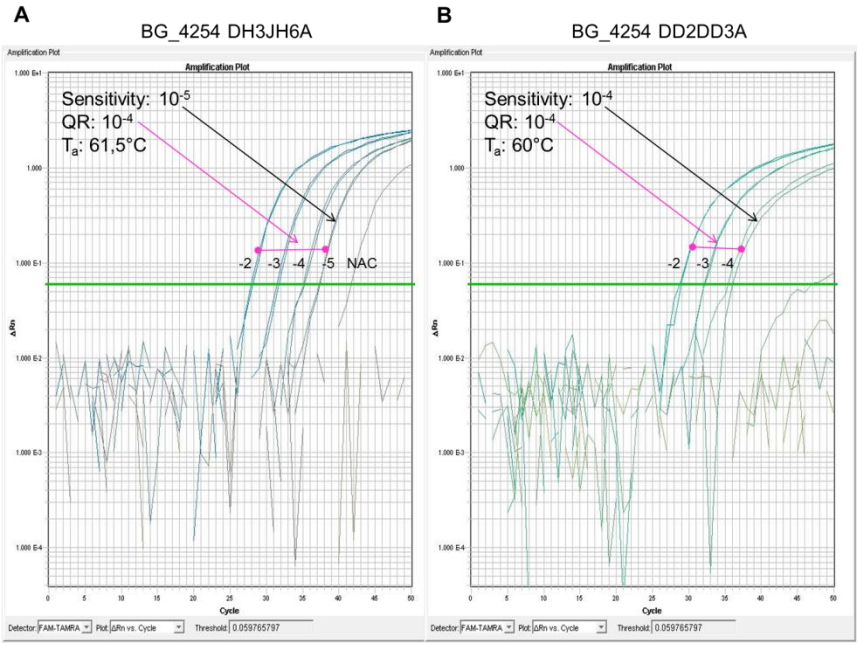

Figure 15. *Retrospective design of Allele-Specific Oligonucleotide (ASO) primers on two rearrangement-specific V-D-J junction regions (DH3JH6, panel A, and DD2D3, panel B) for BG_4254 patient. The sensitivity of ASO-qPCR can reach a detection limit of up to 10^{-5} (1 leukemic cell among 100000 normal cells), indicating the lowest detectable MRD level, while the Quantitative Range (QR) reflects the part of the standard curve in which the MRD level can be quantified reproducibly and accurately. No Amplification Control (NAC) defines the assay background. ASO-qPCR thermal conditions were modified varying the annealing temperature (Ta) to increase sensitivity and specificity.*

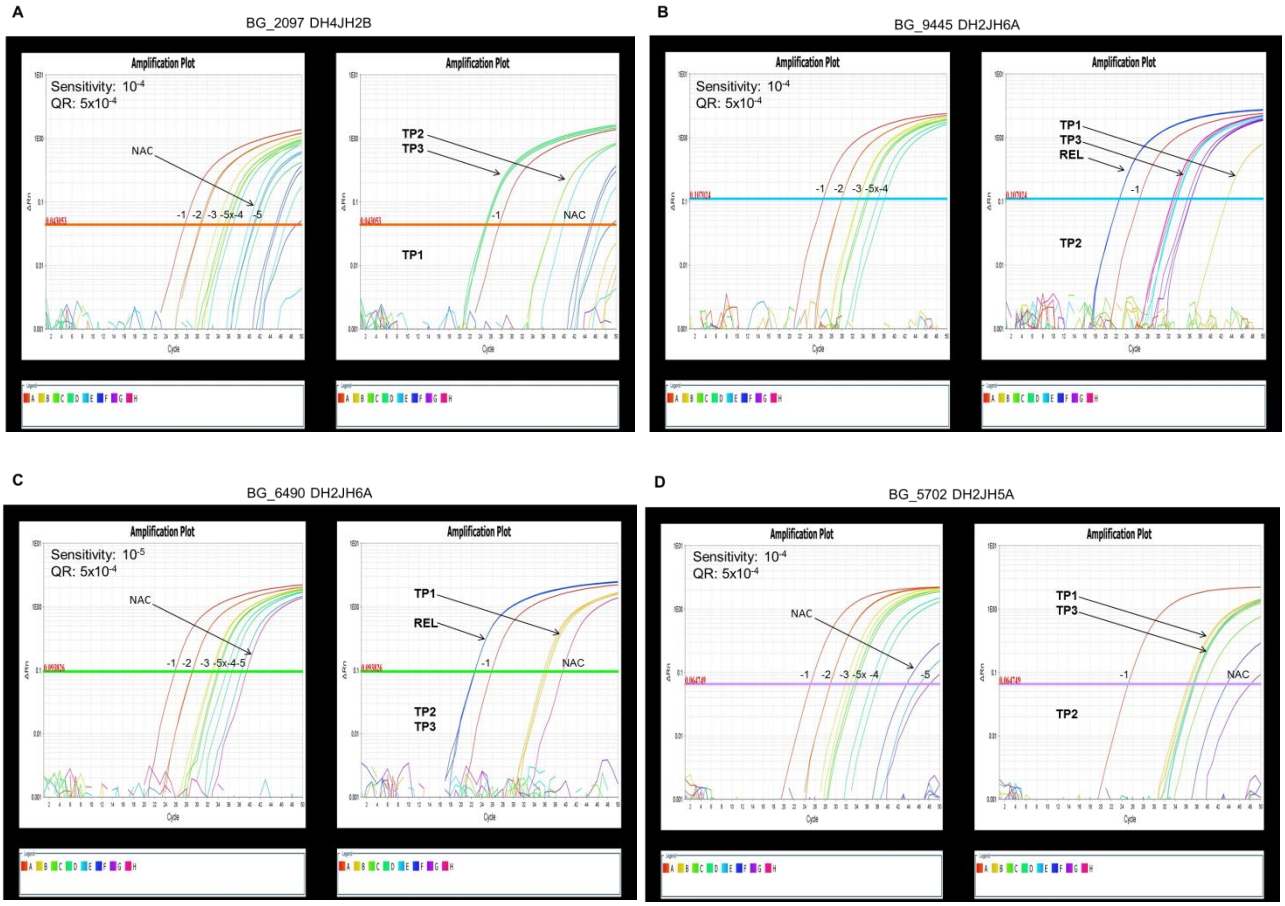

Figure 16. *MRD assessment by ASO-qPCR for BG_2097 (panel A), BG _9445 (panel B), BG_6490 (panel C) and BG_5702 (panel D) patients with available follow-up samples. For each patient, three time-point samples (TP1, TP2, and TP3) and eventually relapse samples (REL) were evaluated. For each ASO-qPCR assay, the patient-specific standard curve made by 10-fold dilutions starting from the patient's sample at diagnosis is on the left panel. The MRD results were expressed as the logarithmic reduction of the leukemic cell burden compared to the diagnosis. For patient BG_9445, TP1 sample showed a faint amplification inhibition (data not shown), thus a weak positivity of the sample cannot be excluded.*

Patient ID	All lineage	Procedure	Probe 1 (P1)	Sensitivity P1	Probe 2 (P2)	Sensitivity P2	TP1	TP2	TP3	Relapse
BG_4205#	B	Na	Na	Na	Na	Na	Na	Na	Na	-
BG_8345*	B	NGS	DH3JH6A	10^{-5}	-	-	Na	Na	Na	-
BG_8345*	B	NGS	DD2DD3A	10^{-4}			Na	Na	Na	-
BG_11584*	B	NGS	Na	Na	Na	Na	Na	Na	Na	-
BG_4254*	B	NGS	DH3JH6A	10^{-5}	-	-	Na	Na	Na	-
BG_4254*	B	NGS	DD2DD3A	10^{-4}			Na	Na	Na	-
BG_10112*	B	NGS	Na	Na	Na	Na	Na	Na	Na	-
BG_10112*	B	NGS	Na	Na	Na	Na	Na	Na	Na	-
BG_2873#	B	Standard	Na	Na	Na	Na	Na	Na	Na	-
BG_2097*	B	NGS	DH4JH2B	10^{-4}	Na	Na	NEG	POS N.Q.	POS (4,7E-01)	-
BG_9445*	B	NGS	DH2JH6	10^{-4}	Na	Na	NEG•	NEG	POS (8,9E-04)	POS (1,5+00)
BG_11543	B	Standard	DD2DD3	10^{-5}	Na	Na	POS (1,0E-03)	POS (1,0E-03)	POS (1,0E-03)	-
BG_6490*	B	NGS	DH2JH6A	10^{-5}	-	-	POS N.Q.	NEG	NEG	POS (1,3+00)
BG_6490*	B	NGS	-	-	DH2JH3A	10-4	POS N.Q.	NEG	NEG	POS (1,1+00)
BG_5702*	B	NGS	DH2JH5A	10^{-4}	Na	Na	POS N.Q.	NEG	POS N.Q.	-
BG_10442	T	Standard	JB2.3A	10^{-5}	Na	Na	NEG	ND	ND	POS (8,1E-01)
BG_8142	T	Standard	JB2.7B	10^{-4}			POS (1,42E-05)	POS (2,3E-02)	Na	-
BG_8646*	T	NGS	DD2DD3B2	10^{-4}	Na	Na	POS (2,0E-01)	POS (2,7E-01)	Na	-
BG_8142	T	Standard	VGIB	10^{-4}	-	-	POS N.Q.	POS (2,5E-02)	Na	-

Table 16. *MRD evaluation of patients enrolled into the NILG -ALL 09/00 clinical trial proved positive for mutations in the TP53 gene. For three patients (BG_11543, BG_10442, and BG_8142) a patient-specific probe based on Ig/TCR rearrangements was designed during the marker screening phase at disease onset, thus allowing the MRD evaluation during the treatment. * For patients indicated with "*", belonging to the cohort of 23 MRD-unknown cases, since effective Ig/TCR molecular markers detection was performed retrospectively by NGS, MRD assessment was performed retrospectively by ASO-qPCR using primers designed based on NGS-derived sequences. # Patients BG_4205 and BG_2873, while MRD-unknown, have not been evaluated by NGS for the retrospective identification of Ig/TCR clonal rearrangements. According to NILG-ALL 09/00 clinical trial, 3 different decisional time-points (TP1, TP2, and TP3) samples were evaluated. NEG indicates MRD-negativity, POS N.Q. indicates a weak, not quantifiable positivity of the sample. For samples MRD-positive (POS), quantifiable residual leukemia level is expressed as logarithmic reduction of the leukemic cells level compared to the diagnosis and indicated in brackets. • For patient BG_9445, TP1 sample showed a faint amplification inhibition, thus a weak positivity of the sample cannot be excluded.*

Na *Not available*
ND *Not Done*

5.6. *Application of capture-based NGS panel on an MRD-unknown prospective cohort allowed to allocate patients to a proper risk-category and consolidation treatment*

The prospective application of our capture-based NGS panel allowed to identify at least one rearrangement in 10 out of 12 newly diagnosed patients (83%) lacking a useful clonal marker by conventional approach, detecting overall 40 clonal rearrangements, located in the following loci: IGH (n=4), IGκ (n=6), IGλ (n=3), TCRβ (n=5), TCRγ (n=14), TCRδ and TCRα (n=8) **(Table 17)**. In 100% of adult B-ALL cases (3 out of 12), we isolated V-D-J or D-J clonal rearrangements, identifying overall 18 rearrangements, including 7 clonal sequences already recognized by Sanger, but not adequate to patient-specific probes generation. The newly identified clones were mainly characterized by uncommon V-D-J combination, missed by standard approach since PCR conventional panel amplifies only the most common rearrangements or some lineage peculiar rearrangements. In particular, an IGκ rearrangement, routinely tested exclusively in B-lineage cases, was recognized as clonal in a T-LL patient, although with low abundance. Furthermore, the NGS approach allowed to detect one TCRγ clonal rearrangement, not previously identified because it was a small subclone (below heteroduplex assay sensitivity) and, interestingly, one IGH and one IGκ clonal rearrangements with high clone abundance and the monoclonal feature. Additionally, two newly identified IGH bi-clonal rearrangements were not tested by standard procedure since the paucity of the diagnostic sample prevented the application of the high-demanding material complete conventional PCR panel **(Figure 17)**. Among the adult T-lineage cases, we identified V-D-J or D-J clonal rearrangements in 7 out of 9 cases. Also in these cases, the failure of the conventional approach relied on the presence of uncommon V-D-J combinations, as well as minor clones and oligoclonalities issue, as previously described for B-ALL patients.

Based on sequences provided by NGS, we designed ASO-qPCR assays for the MRD assessment in those cases for which at least one Ig/TCR clonal marker was identified. In 5 out of 10 cases we obtained at least one patient-specific probe with the required sensitivity of 10^{-5} and QR of 10^{-4}, confirming that the found rearrangements represented major leukemia clones missed by the conventional approach at diagnosis. For the remaining 4 patients it was possible to identify a less sensitive patient-specific probe. In particular, in 2 cases the designed probes reached a sensitivity of 10^{-4} and QR of 10^{-3}, while in the other 2 cases the sensitivity was less than 10^{-4}. Finally, in one case we could not perform the set-up of the ASO-qPCR assay and MRD assessment because of the exhaustion of the diagnostic specimen.

For two patients we used the designed probes to proceed with the MRD assessment on the follow-up samples, thus ensuring the correct allocation of patients in the most adequate risk category. Indeed, in one of these cases, quantifiable positivity of the decisional time points samples during the induction-phase oriented the therapy in favor of the HSCT.

Patient ID	ALL lineage	Locus	V gene	del V	N	del D	D gene	del D	N	del J	J gene	Known	Rearrangement feature
BG_41182	B	TRG	TRGV9*01	-1	0	NA	NA	NA	NA	-8	TRGJ1*02	Y	
BG_41182	B	TRG	TRGV9*01	-2	3	NA	NA	NA	NA	0	TRGJ2*01	Y	
BG_41408	T	TRB	TRBV6-1*01	-10	26	NA	NA	NA	NA	0	TRBJ2-1*01	Y	
BG_41408	T	TRG	TRGV11*01	-2	8	NA	NA	NA	NA	0	TRGJP2*01	Y	
BG_41408	T	TRG	TRGV2*01	-25	0	NA	NA	NA	NA	0	TRGJP2*01	Y	
BG_41408	T	TRD	TRDV2*02	-4	31	NA	NA	NA	NA	0	TRDJ1*01	Y	
BG_41985	B	TRG	TRGV2*02	-4	5	NA	NA	NA	NA	-7	TRGJ1*02	Y	
BG_41985	B	TRG	TRGV3*01	-1	5	NA	NA	NA	NA	0	TRGJP2*01	Y	
BG_41985	B	TRG	TRGV10*02	-1	6	NA	NA	NA	NA	-8	TRGJ1*02	Y	
BG_42228	T	TRB+	NA	NA	NA	NA	TRBD2*01	-2	1	-5	TRBJ2-7*01	Y	
BG_42228	T	TRG	TRGV2*02	0	4	NA	NA	NA	NA	-2	TRGJ1*02	Y	
BG_42228	T	TRG	TRGV2*01	0	6	NA	NA	NA	NA	-9	TRGJ1*02	Y	
BG_42309	B	IGH	IGHV3-23*01	-4	19	-9	IGHD3/OR15-3a*01	-9	8	-11	IGHJ6*02	Y	
BG_42309	B	IGK+	Intron	0	2	NA	NA	NA	NA	-5	KDE	Y	
BG_41985	B	IGH+	NA	NA	NA	NA	IGHD1-7*01	0	1	-2	IGHJ6*02	ND•	
BG_41985	B	IGH+	NA	NA	NA	NA	IGHD1-7*01	-3	11	-17	IGHJ6*02	ND•	
BG_41165	T	TRB	TRBV30*02	-2	10	NA	NA	NA	NA	0	TRBJ1-4*01	Na	Uncommon V/DJ combinations and low represented clone
BG_41165	T	TRA	TRAV19*01	-2	4	NA	NA	NA	NA	-3	TRAJ50*01	Na	Uncommon V/DJ combinations and

													low represented clone
BG_41408	T	TRD	TRDV1*01	-8	20	0	TRDD3*01	-3	5	-3	TRDJ4*01	Na	Uncommon V/DJ combinations
BG_42228	T	TRD	TRDV3*01	-1	9	0	TRDD3*01	-1	0	-3	TRDJ1*01	Na	Uncommon V/DJ combinations
BG_42228	T	TRA	TRAV21*01	-4	1	NA	NA	NA	NA	-1	TRAJ29*01	Na	Uncommon V/DJ combinations
BG_42309	B	IGK	GKV3-15*01	0	3	NA	NA	NA	NA	-1	IGKJ2*03	Na	Uncommon V/DJ combinations
BG_42309	B	IGL	IGLV2-23*01	0	1	NA	NA	NA	NA	0	IGLJ3*02	Na	Uncommon V/DJ combinations
BG_37265	T-LL	IGK+	IGKV3-15*01	-2	5	NA	NA	NA	NA	6	KDE	Na	Uncommon and NT in T-lineage
BG_37265	T-LL	IGK+	IGKV2-28*01	-3	0	NA	NA	NA	NA	0	IGKJ4*01	Na	Uncommon and NT in T-lineage
BG_37265	T-LL	IGK+	IGKV1-33*01	-3	0	NA	NA	NA	NA	0	IGKJ4*01	Na	Uncommon and NT in T-lineage
BG_41182	B	TRA+D	NA	NA	NA	NA	TRDD2*01	-3	7	-3	TRAJ48*01	N - Na	Uncommon V/DJ combinations
BG_41182	B	IGL	IGLV4-3*01	0	2	NA	NA	NA	NA	-2	IGLJ3*02	N	Uncommon V/DJ combinations
BG_41182	B	IGL	IGLV3-1*01	-4	3	NA	NA	NA	NA	-5	IGLJ2*01	N	Uncommon V/DJ combinations
BG_41182	B	TRG	TRGV10*02	-19	8	NA	NA	NA	NA	-17	TRGJP1*01	N	Uncommon V/DJ combinations
BG_41408	T	TRB	TRBV9*01	-4	15	NA	NA	NA	NA	0	TRBJ2-1*01	N	Oligoclonality
BG_40129	T	TRD+	TRDV2*01	0	2	NA	NA	NA	NA	0	TRDD3*01	N	Missed by heteroduplex
BG_41182	B	IGH	IGHV3-38-3*01	11	9	NA	NA	NA	NA	2	IGHJ6*03	N	Missed by heteroduplex
BG_41209	T	TRD+	NA	NA	NA	NA	TRDD2*01	-7	4	-13	TRDD3*01	N	Missed by heteroduplex
BG_42228	T	TRB	TRBV4-1*02	0	4	0	TRBD2*01	-2	0	-5	TRBJ2-1*01	N	Missed by heteroduplex
BG_42309	B	IGK+	IGKV1-17*01	-1	0	NA	NA	NA	NA	8	KDE	N	Missed by heteroduplex
BG_39652	T-LL	TRG	TRGV4*01	-2	0	NA	NA	NA	NA	-5	TRGJ1*01	N	Low represented clone
BG_41165	T	TRG	TRGV4*01	0	4	NA	NA	NA	NA	-11	TRGJ1*02	N	Low represented clone
BG_41165	T	TRG	TRGV11*01	-11	17	NA	NA	NA	NA	6	TRGJP*01	N	Low represented clone

BG_41182	B	TRG	TRGV2*01	-3	2	NA	NA	NA	NA	-8	TRGJ1*02	N	Low represented clone
BG_41733	T	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na	N	
BG_9541	T	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na	N	

Table 17. *Ig/TCR rearrangements identified by NGS in the MRD-unknown cohort of newly diagnosed adult ALL patients. These rearrangements have not been previously identified by standard clonality assessment and Sanger sequencing following the conventional EuroMRD guidelines, or they were not adequate to generate a suitable patient-specific molecular probe for the MRD evaluation. IGK indicates IGκ rearrangements, IGL indicates IGλ rearrangements, TRB indicates TCRβ rearrangements, TRG indicates TCRγ rearrangements, TRD indicates TCRδ rearrangements and TRA indicates TCRα rearrangements. Symbol "+" refers to incomplete rearrangements. For patients BG_41733 and BG_9541, no Ig/TCR rearrangements were detected.*

del V *Nucleotides deletion in V*
N *Number of nucleotides inserted*
del D *Nucleotides deletion in D*
del J *Nucleotide deletion in J*
NA *Not applicable*
Y *Rearrangement already identified by standard procedure*
N *Rearrangement identified by NGS, not previously identified by standard procedure*
Na *Not available*
ND *Not Done*
• *Exhausted diagnostic material*

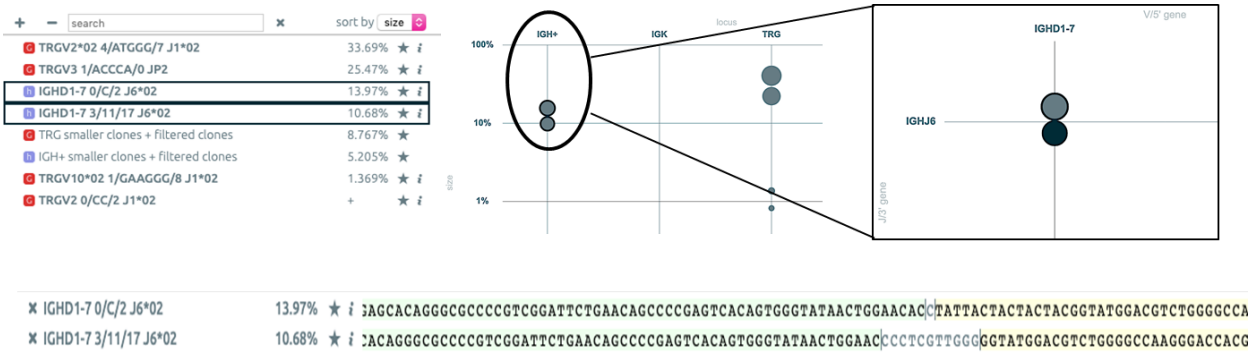

Figure 17. *Newly identified IGH bi-clonal rearrangements for patient BG_41985, not previously tested by standard procedure due to the paucity of the diagnostic sample. The NGS approach allowed the identification of distinct clones and the design of a sensitive ASO-qPCR assay, although the exhaustion of the diagnostic specimen did not permit the MRD monitoring and the follow-up evaluation of the patient.*

5.7. *Concomitant gene variants analysis by NGS contribute to define leukemia genetic profile and to identify prognostic markers in ALL patients.*

In the last years, a number of gene mutations were suggested to be involved in leukemia onset and evolution. However, the incidence of each gene mutation is low in the leukemia population, therefore the single gene mutation analysis approach is ineffective and time-consuming. The inclusion of gene capture and mutation analysis within the Ig/TCR identification NGS approach could be of interest to approach this issue. We modified the Ig/TCR capture assay adding 25 genes following gene of interest (see the methods section).

The Single Nucleotide Variant (SNV) and Insertions/Deletions (Indels) analysis in these genes allowed the identification of 51 gene variants in 18 out of 19 (94.7%) analyzed patients (5 B- and 14 T-lineage), involving *CREBBP*, *EZH2*, *FBXW7*, *FLT3*, *IDH2*, *IKZF1*, *JAK1*, *JAK2*, *JAK3*, *KRAS*, *NOTCH1*, *NRAS*, *PAX5*, *PTEN*, *SH2B3*, *TET2* and *TP53* genes. The Variant Allele Fraction (VAF) of the identified variants, reported in **Table 18**, ranged from 5% to 79%. In all the B-ALL patients we identified a single variant involving one single gene: *CREBBP*, *FLT3*, *JAK2*, *PAX5*, and *TP53*. Also one T-ALL patient harbored only one gene mutated (PTEN) but, in this case, the gene showed two mutations, probably located on different alleles **(Tables 18 and 19)**. In the remaining 13 T-lineage samples we observed the presence of gene variants involving at least two genes and we found a statistically significant correlation between T-lineage and the increase of mutations number per patient (p=0.0282). Concordantly with the literature, variants in *NOTCH1* were present in 7 out of 14 (50%) T-lineage patients of the studied cohort[68] and were associated with gene variants in *NRAS*, *TP53* or *JAK1* and *JAK*. Furthermore, we observed a co-occurrence of variants in *JAK1*, *JAK*3 and *NOTCH1* genes in 3 different patients. In these cases, the *JAK1* gene harbored variants in the pseudokinase domain, while *JAK3* alterations, co-mutated with *JAK1* and *NOTCH*, involved the JH1 kinase domain (p.L857P and p.D846N) or the linker region between the SH2 domain and the pseudokinase domain (p.M511I). This latter variant was previously described in T-cell malignancies as an activating amino acid change.[69] The *NOTCH1* variants associated with *JAK1* and *JAK3* mutations were located in the juxtamembrane heterodimerization (HD) N-terminal domain, but in one of these cases (BG_40129), we identified two further variants of *NOTCH1* (p.P2514H and p.A2463PfsTer14) involving PEST and TAD domain, respectively.

A statistical analysis performed on the co-occurrence of gene variants, although the limited number of patients molecularly profiled, demonstrated a statistically significant association between *JAK1* and *JAK3* ($p<0.0001$) and *JAK1* and *NOTCH1* ($p=0.011$) gene mutations. Moreover, in all the T-LL studied samples we identified the association of *IDH2* mutations and the presence of variants involving *SH2B3* ($p=0.036$), while the co-occurrence of *FLT3* and *IKZF*1 variants was observed in 2 out of 3 T-ALL cases. We did not observe any association between gene variants and specific Ig/TCR rearrangements, as previously described for incomplete IGH rearrangements and *TP53* mutations in B-ALL patients. Of note, different variants (ranging from 2 to 5) were identified in 3 patients in which no Ig/TCR

rearrangements were recognized, thus the risk category of the patients have not been molecularly determined. In those cases, the presence of druggable mutation such as *IDH2* p.R140Q can be useful for therapeutic decisions, while the prognostic value of this mutation is still a matter of debate in several hematologic malignancies, especially in Acute Myeloid Leukemia.[70–72]

Patient ID	Gene	HGVSc	HGVSp	VAF	Read Depth	Alt Read Depth	Consequence
BG_42309	*CREBBP*	NM_004380.2:c.4427C>T	NP_004371.2:p.Pro1476Leu	78,6	641	504	missense_variant
BG_40129	*EZH2*	NM_004456.4:c.1613C>T	NP_004447.2:p.Ser538Leu	51,7	410	212	missense_variant
BG_40129	*EZH2*	NM_004456.4:c.347T>C	NP_004447.2:p.Leu116Pro	43,3	282	122	missense_variant
BG_41733	*EZH2*	NM_004456.4:c.1987T>A	NP_004447.2:p.Tyr663Asn	18,2	292	53	missense_variant
BG_40129	*FBXW7*	NM_033632.3:c.1513C>T	NP_361014.1:p.Arg505Cys	7,4	542	40	missense_variant
BG_4379	*FBXW7*	NM_033632.3:c.62G>A	NP_361014.1:p.Gly21Asp	47,7	2514	1200	missense_variant
BG_41165	*FLT3*	NM_004119.2:c.1779_1793dupTTTCAGAGAATATGA	NP_004110.2:p.Asp593_Tyr597dup	11,8	330	39	inframe_insertion
BG_41985	*FLT3*	NM_004119.2:c.2503G>A	NP_004110.2:p.Asp835Asn	27,6	181	50	missense_variant
BG_4379	*FLT3*	NM_004119.2:c.2864A>G	NP_004110.2:p.Tyr955Cys	51,7	1412	730	missense_variant
BG_37265	*IDH2*	NM_002168.2:c.419G>A	NP_002159.2:p.Arg140Gln	22,7	1122	255	missense_variant
BG_39652	*IDH2*	NM_002168.2:c.547delGACinsAAG	NP_002159.2:p.Asp183Lys	5,6	144	8	missense_variant
BG_41733	*IDH2*	NM_002168.2:c.419G>A	NP_002159.2:p.Arg140Gln	45,8	690	316	missense_variant
BG_4379	*IDH2*	NM_002168.2:c.419G>A	NP_002159.2:p.Arg140Gln	42,9	1559	668	missense_variant
BG_41165	*IKZF1*	NM_006060.4_dupl12.1:c.849G>T	NP_006051.1_dupl12.1:p.Arg284Leu	43,6	165	72	missense_variant
BG_4379	*IKZF1*	NM_006060.4_dupl12.1:c.396G>T	NP_006051.1_dupl12.1:p.Gly133Val	49,0	514	252	missense_variant
BG_40129	*JAK1*	NM_002227.2:c.1954T>C	NP_002218.2:p.Tyr652His	52,5	179	94	missense_variant
BG_42228	*JAK1*	NM_002227.2:c.2107A>T	NP_002218.2:p.Ser703Cys	42,8	566	242	missense_variant
BG_4255	*JAK1*	NM_002227.2:c.2170C>T	NP_002218.2:p.Arg724Cys	12,2	1199	146	missense_variant
BG_41182	*JAK2*	NM_004972.3:c.2171T>C	NP_004963.1:p.Ile724Thr	12,0	465	56	missense_variant

BG_11269	*JAK3*	NM_000215.3:c.1370G>A	NP_000206.2:p.Cys457Tyr	49,2	177	87	missense_variant
BG_40129	*JAK3*	NM_000215.3:c.2570T>C	NP_000206.2:p.Leu857Pro	47,2	301	142	missense_variant
BG_40129	*JAK3*	NM_000215.3:c.2536G>A	NP_000206.2:p.Asp846Asn	13,4	307	41	missense_variant
BG_42228	*JAK3*	NM_000215.3:c.1533G>A	NP_000206.2:p.Met511Ile	44,0	650	286	missense_variant
BG_4255	*JAK3*	NM_000215.3:c.1533G>A	NP_000206.2:p.Met511Ile	29,1	955	278	missense_variant
BG_11269	*KRAS*	NM_033360.2:c.182A>T	NP_203524.1:p.Gln61Leu	6,6	802	53	missense_variant
BG_10442	*NOTCH1*	NM_017617.3:c.7324_7325insTC	NP_060087.3:p.Asp2442ValfsTer36	31,3	412	129	frameshift_variant
BG_37265	*NOTCH1*	NM_017617.3:c.5033T>C	NP_060087.3:p.Leu1678Pro	5,8	291	17	missense_variant
BG_40129	*NOTCH1*	NM_017617.3:c.4799T>A	NP_060087.3:p.Leu1600Gln	43,8	105	46	missense_variant
BG_40129	*NOTCH1*	NM_017617.3:c.7541C>A	NP_060087.3:p.Pro2514His	28,6	276	79	missense_variant
BG_40129	*NOTCH1*	NM_017617.3:c.7387delG	NP_060087.3:p.Ala2463ProfsTer14	5,8	223	13	frameshift_variant
BG_41209	*NOTCH1*	NM_017617.3:c.5165A>C	NP_060087.3:p.Gln1722Pro	24,7	174	43	missense_variant,splice_region_variant
BG_41209	*NOTCH1*	NM_017617.3:c.4787T>C	NP_060087.3:p.Leu1596Pro	6,0	117	7	missense_variant
BG_41408	*NOTCH1*	NM_017617.3:c.3394C>T	NP_060087.3:p.Arg1132Cys	20,6	194	40	missense_variant
BG_42228	*NOTCH1*	NM_017617.3:c.4778T>C	NP_060087.3:p.Leu1593Pro	48,7	189	92	missense_variant
BG_4255	*NOTCH1*	NM_017617.3:c.4776_4777insAGAACC	NP_060087.3:p.Phe1592_Leu1593insArgThr	19,6	168	33	inframe_insertion
BG_4255	*NOTCH1*	NM_017617.3:c.141-4A>G		5,2	155	8	splice_region_variant,intron_variant
BG_37265	*NRAS*	NM_002524.4:c.37G>C	NP_002515.1:p.Gly13Arg	19,3	1464	283	missense_variant
BG_41408	*NRAS*	NM_002524.4:c.35G>A	NP_002515.1:p.Gly12Asp	26,8	1131	303	missense_variant
BG_4379	*NRAS*	NM_002524.4:c.35G>A	NP_002515.1:p.Gly12Asp	43,4	2130	924	missense_variant
BG_5038	*PAX5*	NM_016734.2:c.780+5G>T		45,5	321	146	splice_region_variant,intron_variant

BG_21292	*PTEN*	NM_000314.4:c.493G>T	NP_000305.3:p.Gly165Ter	32,9	347	114	stop_gained,splice_region_variant
BG_21292	*PTEN*	NM_000314.4:c.736_737insAG	NP_000305.3:p.Pro246GlnfsTer11	34,3	1130	388	frameshift_variant
BG_37265	*SH2B3*	NM_005475.2:c.1345G>A	NP_005466.1:p.Glu449Lys	5,8	924	54	missense_variant
BG_39652	*SH2B3*	NM_005475.2:c.927-2delAGinsCT		5,4	112	6	splice_acceptor_variant
BG_4255	*SH2B3*	NM_005475.2:c.1038dupG	NP_005466.1:p.Leu347AlafsTer38	50,2	944	474	frameshift_variant
BG_39652	*TET2*	NM_001127208.2:c.1588delCAinsTG	NP_001120680.1:p.Gln530Trp	6,2	242	15	stop_gained
BG_41165	*TET2*	NM_001127208.2:c.5733delA	NP_001120680.1:p.Lys1911AsnfsTer39	35,7	493	176	frameshift_variant
BG_10442	*TP53*	NM_000546.5:c.684_685insGGGGTTTGACC	NP_000537.3:p.Cys229GlyfsTer3	5,9	236	14	stop_gained,frameshift_variant
BG_10442	*TP53*	NM_000546.5:c.651_654dupGGTG	NP_000537.3:p.Pro219GlyfsTer4	33,7	943	318	frameshift_variant
BG_11584	*TP53*	NM_000546.5:c.844C>G	NP_000537.3:p.Arg282Gly	4,6	1034	47	missense_variant
BG_41209	*TP53*	NM_000546.5:c.390_426delCAACAAGATGTTTTGCCAACTGGCCAAGACCTGCCCT	NP_000537.3:p.Asn131CysfsTer27	35,0	117	41	frameshift_variant

Table 18. *List of gene variants identified in 18 out of 19 adult ALL patients of which 7 formerly evaluated for Ig/TCR rearrangements and 12 newly diagnosed ALL patients lacking molecular Ig/TCR probe.*

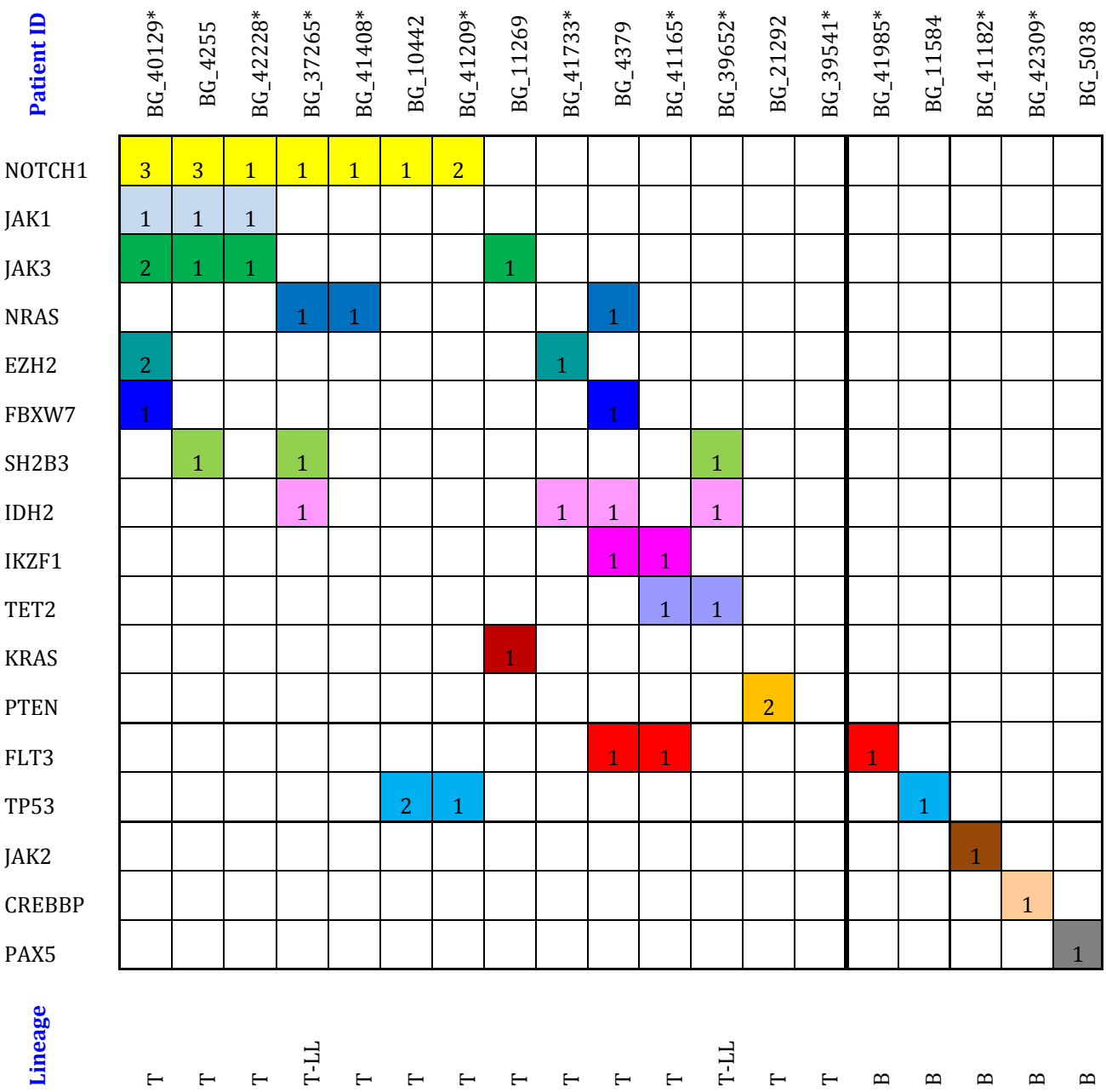

Table 19. *Summary of gene variants identified in 18 out of 19 adult ALL patients. The filled boxes relate to the presence and the number of gene variants identified by capture-NGS. The 12 newly diagnosed ALL patients are indicated with "*". For patient BG_39541 no gene variants nor Ig/TCR rearrangements were detected.*

6. DISCUSSION

The identification of the leukemia-associated clonal Ig/TCR is instrumental for Minimal Residual Disease (MRD) monitoring during disease treatment in the Acute Lymphoblastic Leukemia context. Despite this process has been developed and ruled during the last 20 years within European collaborative networks (EuroClonality and EuroMRD, part of ESHLO foundation) it is laborious, time-consuming and, in some cases, ineffective. Therefore, MRD stratification using the standard ASO-qPCR approach is nowadays feasible in about 90% of cases.[30] With the aim to develop a simplified and possibly more efficient method, in this work, we designed a novel capture-based NGS panel useful for the Ig/TCR-based clonal markers screening at diagnosis of ALL.

The NGS capture design was tested on 10 diagnostic materials derived from patients previously studied for rearrangements' identification. This system identified all the known Ig/TCR clonal rearrangements, so demonstrating to be a powerful tool for the identification of clonalities at disease presentation. Therefore, our NGS approach resulted in being a valid alternative to standard techniques, being less laborious and easier to perform by the same well-trained personnel. Furthermore, NGS proved faster compared to multistep standard workflow, allowing an efficient patient-specific probe design for timely MRD assessment, based on NGS-derived sequences. Indeed, the MRD persistence during therapy reflects the patients' limited chemosensitivity, whereas patients responsive to the firsts cycles of treatment are less prone to develop leukemia relapse. In childhood protocols, early MRD evaluation (day 15 and 33 of treatment) by flow cytometry allowed to identify patients with a rapid leukemia clearance that correlated to better outcome.[73] Conversely, patients with persistence of leukemia at these early time points had a poorer outcome. Flow cytometric MRD evaluation is faster than molecular procedure, but the sensitivity is less profound and data interpretation is still operator-dependent. However, conventional molecular methods do not yet allow fast MRD evaluation. NGS-based approaches are facing this issue and demonstrated effective in shortening the time to results, but they are not yet ready to be as fast as needed. Our approach allowed to simplify the first phase of clonal rearrangements detection and represents a step toward the final goal.

Our capture-based NGS assay not only identified all clonal rearrangements previously detected by standard assessment but allowed the identification of additional clonalities, not previously recognized and later confirmed, providing supplementary information on biclonal, oligoclonal and uncommon rearrangements, not isolated by conventional methods. This aspect is particularly important in patients in which no Ig/TCR clonalities can be identified with standard procedure. In fact, in our clinical study, this cohort (defined as MRD-unknown, MRD-unk) represented 16% of the enrolled patients.[15] These patients could not be followed during the treatment with an MRD specific probe,

consenting a treatment only based on clinical features that, often, proved not completely adequate to define the relapse risk of the patient. Therefore, we verify this new power of the assay on patients prospectively enrolled into the active trial in which the standard method fails. In this cohort of patients, 5 out of 10 (50%) were recovered and a useful MRD assay could have been designed on clonal sequences provided by NGS capture assay, so allowing a proper risk-class assignment.

Moreover, NGS offered an important improvement in terms of increased sensitivity resulting in the possibility to identify minor leukemia clones which could eventually be responsible for unexpected relapse. Indeed, monitoring MRD by ASO-qPCR requires to choose one or two clonal rearrangements for follow-up measurements, usually selected for better sensitivity. This often reflects also the abundance of the leukemic clone among the cell population. Therefore, the ASO-qPCR approach necessarily leads to losing track of minor subclones that might resist to therapy and be responsible for disease resistance or reappearance. In view of this and considering our results, it could be desirable to explore all the possible rearrangements, also the uncommon ones not included into the standard PCR panel or amplicon-based NGS panel. Despite the presence of specific and different surface receptors on B- and T-lineage ALL, the rearrangements process precedes the lineage commitment. Therefore, even in B-precursor ALL, some incomplete or complete non-functional T-cell receptor rearrangements are present, as well as Ig combination can be found in T-ALL. For this reason, some but not all lineage-specific rearrangements are investigated either in T- and B-lineage ALL with current methods. Indeed, it could be suitable to test all possible rearrangements in any ALL, in order to avoid the missing of unexpected clonal markers. That was possible with our capture-based NGS. This assay allowed to identify one T-LL patient (BG_37265) positive for IGKV3-15*01/KDE rearrangement, not previously identified because this rearrangement is routinely tested exclusively in B-lineage ALL.

In one our previous work we identified *TP53* mutations in 8% of patients enrolled into the NILG-ALL 09/00 clinical trial, which reached clinical remission but suffered a relapse in a very short time post CR.[61] Interestingly enough, in the 64% of mutated patients no informative Ig/TCR rearrangements were identified for the design of a suitable patient-specific molecular probe at time of marker screening. Based on this result, it has been hypothesized that *TP53* mutations occurred in immature, not yet rearranged leukemia blasts, maintaining intrinsic higher staminality and chemoresistance features, responsible for the leukemia clone persistence and relapse. However, in the MRD-unknown B-ALL patients carrying *TP53* mutations included in our analysis, we not only retrospectively identified at least one Ig/TCR rearrangement, but we also noted a higher prevalence of incomplete IGH rearrangements, compared to the *TP53* wild type counterpart. Since the presence of *TP53* gene mutations increases the predisposition to relapse, the definition of the mutational status of this gene at diagnosis must be included in the diagnostic workup of adult B-ALL, especially in presence of associated IGH incomplete rearrangements. In view of all these data, the risk-category definition and

the MRD-driven treatment could benefit from both *TP53* mutational status analysis at diagnosis and MRD evaluation during early treatment.

This observation and the demanding process to obtain *TP53* mutation analysis prompted us to include *TP53* and a number of prognostically significant genes into the capture design. This resulted in the identification of mutations in ALL cases. In particular, all the B-ALL cases tested to investigate the mutational status of ALL-related genes of interest, harbored single gene variants involving *CREBBP*, *FLT3*, *JAK2*, *PAX5*, and *TP53.* On the contrary, in all the T-ALL derived samples we observed the presence of gene variants involving at least two genes and we also demonstrated the increase of the mutations number per T-lineage patient, compared to B-ALL cases. In this subgroup of patients, we did not observe any association between gene variants and specific Ig/TCR rearrangements, as previously described for incomplete IGH rearrangements and *TP53* mutations in B-ALL patients. Based on the data so far obtained, we can conclude that for T-ALL patients the inclusion of mutational status analysis of leukemia-associated genes could improve more and more not only the general knowledge about leukemogenesis, but it could have important consequences in clinical practice. Indeed, we identified several variants (ranging from 2 to 5) in 3 patients in which no Ig/TCR rearrangements were recognized, thus the risk-category of the patients has not been molecularly determined. In those cases, the presence of druggable mutation could be considered for therapeutic decisions for the off-label use of specific targeted drugs, while the prognostic value of the majority of mutations is still a matter of debate in several hematologic malignancies. Of note, even in the absence of druggable mutations, the presence of gene variants could be employed for the setting up of specific assay using these abnormalities as markers to determine the MRD level during the time. For instance, highly sensitive droplet digital PCR (ddPCR) techniques can be a valuable surrogate to ASO-qPCR to describe the variations of a given burden gene variant during the time, associated with disease evolution.

The capture-based NGS approach proved very important also for saving the diagnostic material collected at disease onset. Indeed, the amount of 1 μg of genomic DNA required for the complete NGS panel for clonal markers screening, including the study of rare rearrangements and genetic profiling, is much lower than the amount of DNA needed for the amplicon-based techniques. This consideration is pivotal especially from the perspective in which nowadays the MRD evaluation is still performed by ASO-qPCR, using serial dilutions of the diagnostic specimen, and MRD quantification is expressed as logarithmic reduction compared to the leukemia burden at disease onset. Despite this method proved very sensitive and, in most cases, functional to predict leukemia relapse, some low positive results cannot be distinguished from non-specific signals caused by miss-annealing of Allele-Specific Oligonucleotides to similar sequences from normal leukocytes. Initially, it has been hypothesized that the NGS approach could overcome this limitation since this technology provides single sequences that can be compared to the leukemic clone sequence. Unfortunately, but not surprising, our capture-based

NGS approach proved to be low sensitive in MRD detection, identifying about 1 leukemia cell within 100 normal leukocytes, reaching a sensitivity level insufficient for an appropriate MRD evaluation. Subsequently, due to the intrinsic structure of the capture-based method and the lack of an appropriate laboratories' network focused on the attempt to increase sensitivity, this aim was no subsequent carried out. Moreover, newly described techniques as amplicon NGS or ddPCR, aiming to achieve an absolute quantification of residual leukemia levels, whereas promising, are still under investigation and not yet standardized. Additionally, the saving of diagnostic material is essential in hypocellular specimens or when *punctio sicca* occurs and the follow-up samples analysis is required for a long time.

In our work, all B-ALL MRD-unknown patients tested retrospectively as well prospectively, showed at least one clonal rearrangement, while all patients still lacking molecular marker in our cohort belonged to T-lineage group.

For two of these patients, we can speculate that the tested diagnostic sample was not adequate for the appropriate analysis. In fact, in one case (BG_4005) the blasts percent content in the diagnostic sample was not available and we cannot exclude that the specimen was not leukemia-representative. Likewise, the blasts percentage in the bone marrow diagnostic sample for patient BG_39541 was around 5%, not sufficient to the proper leukemia diagnosis according to the World Health Organization criteria. Even then, we can suppose that the available diagnostic specimen was not correctly evaluable. As a matter of fact, sample from BG_39541 patient has been evaluated both for Ig/TCR rearrangements and gene mutational status analysis, but neither Ig/TCR clonal markers nor gene variants were identified. On the contrary, the diagnostic sample from BG_37265 patient, containing a comparable blasts percentage (5-10%), showed evidence of both Ig/TCR rearrangements and gene variants. Interestingly, we identified for this patient mutations in *IDH2* and *NRAS* genes with Variant Allele Fraction (VAF) around 20% and mutations in *NOTHC1* and *SH2B3* genes with VAF of 6%. Unlike for germline designation of a variant, for which the VAF should be set approximately around 50% for a heterozygous variant and 100% for a homozygous variant, for somatic variants the VAF percentage correlates with the tumor cells content among the complete, heterogeneous cell population.[74] Based on this data, it may be supposed that the leukemic cells in the above-mentioned diagnostic sample rather represented around 20% of global cell population. Actually, the blasts percentage assessment is performed by flow cytometry analysis over the whole peripheral blood or bone marrow specimen, determining a miscalculation and a discrepancy in the determination of blasts percentage content. Indeed, after the separation of mononuclear cells population, leukemic lymphoblasts result enriched in the sample, representing the total of the cell population in further analysis. Based on this assumption, we could assume that variants involving *IDH2* and *NRAS* were present in homozygosity, although *NOTCH1* and *SH2B3* variants were present in heterozygosity.

However, the results so far obtained could suggest the definition of a small subgroup of very immature T-ALL in which Ig/TCR rearrangements not yet occurred, although further analysis is required to speculate the maturation stage of leukemic lymphoblasts and to exclude technical limitations. Nevertheless, our NGS approach allowed to identify very small leukemia clones, representing less than 1% of the global leukemia population, but lacking a spike-in control to normalize clonal rearrangements size compared to physiological rearrangements in normal lymphocytes, is not possible to exclude a technical bias. In any case, it is important to note that the group of prospectively-analyzed MRD-unknown patients is characterized by a strong prevalence of T-lineage cases, while globally the ratio between B- and T-ALL cases is unbalanced towards B-lineage occurrence.

In conclusion, our results indicate that the prospective application of NGS remarkably simplifies the assessment of clonality in adult ALL, provides additional information about the mutational status of some ALL-related genes and it is becoming the new standard to identify the most informative Ig and TCR markers for the prospective evaluation of MRD.

7. BIBLIOGRAPHY

1. Terwilliger T, Abdul-Hay M. Acute lymphoblastic leukemia: a comprehensive review and 2017 update. *Blood Cancer Journal*. 2017;7(6):e577–e577.
2. Hoelzer D, Bassan R, Dombret H, et al. Acute lymphoblastic leukaemia in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology*. 2016;mdw025.
3. Bassan R, Gatta G, Tondini C, Willemze R. Adult acute lymphoblastic leukaemia. *Critical Reviews in Oncology/Hematology*. 2004;50(3):223–261.
4. Vardiman JW, Thiele J, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*. 2009;114(5):937–951.
5. Bennett JM, Catovsky D, Daniel M-T, et al. Proposals for the Classification of the Acute Leukaemias French-American-British (FAB) Co-operative Group. *British Journal of Haematology*. 1976;33(4):451–458.
6. Vardiman JW, Thiele J, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*. 2009;114(5):937–951.
7. Bene MC, Castoldi G, Knapp W, et al. Proposals for the immunological classification of acute leukemias. European Group for the Immunological Characterization of Leukemias (EGIL). *Leukemia*. 1995;9(10):1783–1786.
8. Béné MC, Nebe T, Bettelheim P, et al. Immunophenotyping of acute leukemia and lymphoproliferative disorders: a consensus proposal of the European LeukemiaNet Work Package 10. *Leukemia*. 2011;25(4):567–574.
9. Pui CH, Crist WM, Look AT. Biology and clinical significance of cytogenetic abnormalities in childhood acute lymphoblastic leukemia. *Blood*. 1990;76(8):1449–1463.
10. Jabbour E, O'Brien S, Konopleva M, Kantarjian H. New insights into the pathophysiology and therapy of adult acute lymphoblastic leukemia. *Cancer*. 2015;121(15):2517–2528.
11. Bassan R, Hoelzer D. Modern therapy of acute lymphoblastic leukemia. *J. Clin. Oncol.* 2011;29(5):532–543.
12. Hoelzer D. Treatment of acute lymphoblastic leukemia. *Semin. Hematol.* 1994;31(1):1–15.
13. Gökbuget N, Stanze D, Beck J, et al. Outcome of relapsed adult lymphoblastic leukemia depends on response to salvage chemotherapy, prognostic factors, and performance of stem cell transplantation. *Blood*. 2012;120(10):2032–2041.
14. van Dongen JJM, van der Velden VHJ, Brüggemann M, Orfao A. Minimal residual disease diagnostics in acute lymphoblastic leukemia: need for sensitive, fast, and standardized technologies. *Blood*. 2015;125(26):3996–4009.
15. Bassan R, Spinelli O, Oldani E, et al. Improved risk classification for risk-specific therapy based on the molecular study of minimal residual disease (MRD) in adult acute lymphoblastic leukemia (ALL). *Blood*. 2009;113(18):4153–4162.
16. Bassan R, Rossi G, Pogliani EM, et al. Chemotherapy-Phased Imatinib Pulses Improve Long-Term Outcome of Adult Patients With Philadelphia Chromosome-Positive Acute Lymphoblastic Leukemia: Northern Italy Leukemia Group Protocol 09/00. *Journal of Clinical Oncology*. 2010;
17. Cavé H, van der Werff ten Bosch J, Suciu S, et al. Clinical significance of minimal residual disease in childhood acute lymphoblastic leukemia. European Organization for Research and Treatment of Cancer--Childhood Leukemia Cooperative Group. *N. Engl. J. Med.* 1998;339(9):591–598.
18. Brüggemann M, Raff T, Flohr T, et al. Clinical significance of minimal residual disease quantification in adult patients with standard-risk acute lymphoblastic leukemia. *Blood*. 2006;107(3):1116–1123.
19. Beldjord K, Chevret S, Asnafi V, et al. Oncogenetics and minimal residual disease are independent outcome predictors in adult patients with acute lymphoblastic leukemia. *Blood*. 2014;123(24):3739–3749.
20. Ribera J-M, Oriol A, Morgades M, et al. Treatment of High-Risk Philadelphia Chromosome–Negative Acute Lymphoblastic Leukemia in Adolescents and Adults According to Early Cytologic Response and Minimal Residual Disease After Consolidation Assessed by Flow Cytometry: Final Results of

the PETHEMA ALL-AR-03 Trial. *Journal of Clinical Oncology*. 2014;
21. Brüggemann M, Raff T, Kneba M. Has MRD monitoring superseded other prognostic factors in adult ALL? *Blood*. 2012;120(23):4470–4481.
22. van der Velden VHJ, Hochhaus A, Cazzaniga G, et al. Detection of minimal residual disease in hematologic malignancies by real-time quantitative PCR: principles, approaches, and laboratory aspects. *Leukemia*. 2003;17(6):1013–1034.
23. van der Velden VHJ, Cazzaniga G, Schrauder A, et al. Analysis of minimal residual disease by Ig/TCR gene rearrangements: guidelines for interpretation of real-time quantitative PCR data. *Leukemia*. 2007;21(4):604–611.
24. van Dongen JJM, Lhermitte L, Böttcher S, et al. EuroFlow antibody panels for standardized n-dimensional flow cytometric immunophenotyping of normal, reactive and malignant leukocytes. *Leukemia*. 2012;26(9):1908–1975.
25. van Dongen JJ, Macintyre EA, Gabert JA, et al. Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease. Report of the BIOMED-1 Concerted Action: investigation of minimal residual disease in acute leukemia. *Leukemia*. 1999;13(12):1901–1928.
26. Gabert J, Beillard E, van der Velden VHJ, et al. Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia – A Europe Against Cancer Program. *Leukemia*. 2003;17(12):2318–2357.
27. Campana D. Determination of minimal residual disease in leukaemia patients. *British Journal of Haematology*. 2003;121(6):823–838.
28. Pongers-Willemse MJ, Seriu T, Stolz F, et al. Primers and protocols for standardized detection of minimal residual disease in acute lymphoblastic leukemia using immunoglobulin and T cell receptor gene rearrangements and TAL1 deletions as PCR targets: report of the BIOMED-1 CONCERTED ACTION: investigation of minimal residual disease in acute leukemia. *Leukemia*. 1999;13(1):110–118.
29. van Dongen JJM, Langerak AW, Brüggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia*. 2003;17(12):2257–2317.
30. Brüggemann M, Kotrova M. Minimal residual disease in adult ALL: technical aspects and implications for correct clinical interpretation. *Blood Adv*. 2017;1(25):2456–2466.
31. Charles A Janeway J, Travers P, Walport M, Shlomchik MJ. The generation of diversity in immunoglobulins. *Immunobiology: The Immune System in Health and Disease. 5th edition*. 2001;
32. Calis JJA, Rosenberg BR. Characterizing immune repertoires by high throughput sequencing: strategies and applications. *Trends Immunol.* 2014;35(12):581–590.
33. Langerak AW, van Dongen JJM. Multiple clonal Ig/TCR products: implications for interpretation of clonality findings. *Journal of Hematopathology*. 2012;5(1–2):35–43.
34. Nunes V, Cazzaniga G, Biondi A. An update on PCR use for minimal residual disease monitoring in acute lymphoblastic leukemia. *Expert Rev. Mol. Diagn.* 2017;17(11):953–963.
35. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 1977;74(12):5463–5467.
36. Smith LM, Sanders JZ, Kaiser RJ, et al. Fluorescence detection in automated DNA sequence analysis. *Nature*. 1986;321(6071):674–679.
37. Rizzo JM, Buck MJ. Key Principles and Clinical Applications of "Next-Generation" DNA Sequencing. *Cancer Prev Res*. 2012;5(7):887–900.
38. Metzker ML. Sequencing technologies — the next generation. *Nature Reviews Genetics*. 2010;11(1):31–46.
39. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics*. 2016;17(6):333–351.
40. Park JY, Kricka LJ, Fortina P. Next-generation sequencing in the clinic. *Nature Biotechnology*. 2013;31(11):990–992.
41. Yohe S, Thyagarajan B. Review of Clinical Next-Generation Sequencing. *Archives of Pathology & Laboratory Medicine*. 2017;141(11):1544–1557.

42. Samorodnitsky E, Jewell BM, Hagopian R, et al. Evaluation of Hybridization Capture Versus Amplicon-Based Methods for Whole-Exome Sequencing. *Human Mutation*. 2015;36(9):903–914.
43. Jennings LJ, Arcila ME, Corless C, et al. Guidelines for Validation of Next-Generation Sequencing–Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. *The Journal of Molecular Diagnostics*. 2017;19(3):341–365.
44. Deorowicz S, Grabowski S. Compression of DNA sequence reads in FASTQ format. *Bioinformatics*. 2011;27(6):860–862.
45. Ewing B, Hillier L, Wendl MC, Green P. Base-Calling of Automated Sequencer Traces UsingPhred. I. Accuracy Assessment. *Genome Res.* 1998;8(3):175–185.
46. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–2079.
47. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*. 2011;43(5):491–498.
48. Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia – A Europe Against Cancer Program | Leukemia.
49. Faham M, Zheng J, Moorhead M, et al. Deep-sequencing approach for minimal residual disease detection in acute lymphoblastic leukemia. *Blood*. 2012;120(26):5173–5180.
50. Ladetto M, Brüggemann M, Monitillo L, et al. Next-generation sequencing and real-time quantitative PCR for minimal residual disease detection in B-cell disorders. *Leukemia*. 2014;28(6):1299–1307.
51. Langerak AW, Brüggemann M, Davi F, et al. High-Throughput Immunogenetics for Clinical and Research Applications in Immunohematology: Potential and Challenges. *J. Immunol.* 2017;198(10):3765–3774.
52. Hung SS, Meissner B, Chavez EA, et al. Assessment of Capture and Amplicon-Based Approaches for the Development of a Targeted Next-Generation Sequencing Pipeline to Personalize Lymphoma Management. *The Journal of Molecular Diagnostics*. 2018;20(2):203–214.
53. Wood B, Wu D, Crossley B, et al. Measurable residual disease detection by high-throughput sequencing improves risk stratification for pediatric B-ALL. *Blood*. 2018;131(12):1350–1359.
54. Brüggemann M, Kotrová M, Knecht H, et al. Standardized next-generation sequencing of immunoglobulin and T-cell receptor gene recombinations for MRD marker identification in acute lymphoblastic leukaemia; a EuroClonality-NGS validation study. *Leukemia*. 2019;
55. Knecht H, Reigl T, Kotrová M, et al. Quality control and quantification in IG/TR next-generation sequencing marker identification: protocols and bioinformatic functionalities by EuroClonality-NGS. *Leukemia*. 2019;
56. Sánchez R, Ayala R, Martínez-López J. Minimal Residual Disease Monitoring with Next-Generation Sequencing Methodologies in Hematological Malignancies. *International Journal of Molecular Sciences*. 2019;20(11):2832.
57. Scheijen B, Meijers RWJ, Rijntjes J, et al. Next-generation sequencing of immunoglobulin gene rearrangements for clonality assessment: a technical feasibility study by EuroClonality-NGS. *Leukemia*. 2019;33(9):2227–2240.
58. Sala Torra O, Othus M, Williamson DW, et al. Next-Generation Sequencing in Adult B Cell Acute Lymphoblastic Leukemia Patients. *Biol. Blood Marrow Transplant.* 2017;23(4):691–696.
59. Salson M, Giraud M, Caillault A, et al. High-throughput sequencing in acute lymphoblastic leukemia: Follow-up of minimal residual disease and emergence of new clones. *Leuk. Res.* 2017;53:1–7.
60. Langerak AW, Groenen PJTA, Brüggemann M, et al. EuroClonality/BIOMED-2 guidelines for interpretation and reporting of Ig/TCR clonality testing in suspected lymphoproliferations. *Leukemia*. 2012;26(10):2159–2171.
61. Salmoiraghi S, Montalvo MLG, Ubiali G, et al. Mutations of TP53 gene in adult acute lymphoblastic leukemia at diagnosis do not affect the achievement of hematologic response but correlate with early relapse and very poor survival. *Haematologica*. 2016;101(6):e245–e248.
62. Mullighan CG. The molecular genetic makeup of acute lymphoblastic leukemia. *Hematology Am Soc Hematol Educ Program*. 2012;2012:389–396.

63. Giraud M, Salson M, Duez M, et al. Fast multiclonal clusterization of V(D)J recombinations from high-throughput sequencing. *BMC Genomics*. 2014;15:409.
64. Duez M, Giraud M, Herbert R, et al. Vidjil: A Web Platform for Analysis of High-Throughput Repertoire Sequencing. *PLoS ONE*. 2016;11(11):e0166126.
65. Langerak AW, Brüggemann M, Davi F, et al. High-Throughput Immunogenetics for Clinical and Research Applications in Immunohematology: Potential and Challenges. *The Journal of Immunology*. 2017;198(10):3765–3774.
66. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotech*. 2011;29(1):24–26.
67. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform*. 2013;14(2):178–192.
68. Notch-1 Mutations Are Secondary Events in Some Patients with T-Cell Acute Lymphoblastic Leukemia | Clinical Cancer Research.
69. Degryse S, Bornschein S, de Bock CE, et al. Mutant JAK3 signaling is increased by loss of wild-type JAK3 or by acquisition of secondary JAK3 mutations in T-ALL. *Blood*. 2018;131(4):421–425.
70. Green CL, Evans CM, Zhao L, et al. The prognostic significance of IDH2 mutations in AML depends on the location of the mutation. *Blood*. 2011;118(2):409–412.
71. Zhou K-G, Jiang L-J, Shang Z, et al. Potential application of IDH1 and IDH2 mutations as prognostic indicators in non-promyelocytic acute myeloid leukemia: a meta-analysis. *Leuk. Lymphoma*. 2012;53(12):2423–2429.
72. Patel JP, Gönen M, Figueroa ME, et al. Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *N. Engl. J. Med.* 2012;366(12):1079–1089.
73. Basso G, Veltroni M, Valsecchi MG, et al. Risk of Relapse of Childhood Acute Lymphoblastic Leukemia Is Predicted By Flow Cytometric Measurement of Residual Disease on Day 15 Bone Marrow. *JCO*. 2009;27(31):5168–5174.
74. Li MM, Datto M, Duncavage EJ, et al. Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer. *The Journal of Molecular Diagnostics*. 2017;19(1):4–23.

8. SITOGRAPHY

https://clinicaltrials.gov/

http://www.vidjil.org/

http://software.broadinstitute.org/software/igv/

https://www.ncbi.nlm.nih.gov/clinvar/

https://cancer.sanger.ac.uk/cosmic

http://exac.broadinstitute.org/

https://gnomad.broadinstitute.org/

https://www.ncbi.nlm.nih.gov/snp/

http://www.imgt.org/

http://www.eslho.org/usr/index.php

http://www.euroclonality.org/

http://www.euromrd.org/usr/pub/pub.php

9. SCIENTIFIC PRODUCTS

Abstracts:

• C. Caprioli, T. Intermesoli, O. Spinelli, S. Salmoiraghi, P. Zanghì, **R. Cavagna**, A. Michelato, K. Buklijas, F. Delaini, E. Oldani, G. Gianfaldoni, D. Ferrero, E. Terruzzi, L. De Paoli, C. Cattaneo, E. Borlenghi, I. Cavattoni, M. Tajana, A.M. Scattolin, D.G. Mattei, P. Corradini, L. Campiotti, F. Ciceri, M. Bernardi, E. Todisco, A. Cortelezzi, S. Cortelazzo, E. Audisio, F. Marmont, A. Bosi, B. Falini, C. Pavoni, R. Bassan, A. Rambaldi and F. Lussana. *Clinical Impact of an Accurate Genetic Characterization of Older Acute Myeloid Leukemia Patients: A Report from the Northern Italy Leukemia Group (NILG) Randomized Trial 02/06.* Blood. 2018;132(Suppl 1):434–434.

• **R. Cavagna**, M.L. Guinea Montalvo, M. Tosi, M. Paris, A. Rambaldi, A. Mosca, O. Spinelli. *Next Generation Sequencing improves the identification of Immunoglobulin/T-cell receptors clonal markers in Acute Lymphoblastic Leukemia.* 2nd European PhD and Postdoc symposium: The Promise of Future Medicine: From Research to Therapy, Copenhagen, November 2018. Award: Travel grant.

• **R. Cavagna**, A. Pansa, S. Salmoiraghi, K. Buklijas, A. Michelato, L. Zannino, T. Intermesoli, F. Lussana, C. Caprioli, P. Stefanoni, G. Cassina, B. Facchinetti, A. Mosca, A. Rambaldi, U. Giussani, O. Spinelli. *The integration of multiple genetic analysis improves the diagnostic process and provides a better risk-stratification in Acute Myeloid Leukemia.* Foundation FS. XV Congress of the Italian Society of Experimental Hematology Rimini, Italy, October 18-20, 2018. *Haematologica*. 2018;103(s3):S1–S129.

• O. Spinelli, S. Salmoiraghi, P. Zanghì, **R. Cavagna**, A. Michelato, K. Buklijas, L. Zannino, T. Intermesoli, F. Lussana, F. Delaini, E. Oldani, C. Caprioli, P. Stefanoni, G. Gianfaldoni, F. Marmont, D. Ferrero, E. Terruzzi, L. De Paoli, G. Rossi, E. Borlenghi, I. Cavattoni, M. Tajana, A.M. Scattolin, D. Mattei, P. Corradini, L. Campiotti, F. Ciceri, M. Bernardi, E. Todisco, A. Cortelezzi, S. Cortelazzo, E. Audisio, A. Bosi, B. Falini, C. Pavoni, R. Bassan, A. Rambaldi. *Molecular profile by Next Generation Sequencing of Acute Myeloid Leukemia with normal karyotype: clinical results from the prospective trial 02/06 of the Northern Italy Leukemia Group (NILG).* 23rd Congress of the European Hematology Association Stockholm, Sweden, June 14-17, 2018. HemaSphere. 2018;2(S1):1. Foundation FS. XV Congress of the Italian Society of Experimental Hematology Rimini, Italy, October 18-20, 2018. *Haematologica*. 2018;103(s3):S1–S129.

• Pansa, S. Salmoiraghi, **R. Cavagna**, K. Buklijas, A. Michelato, L. Zannino, G. Cassina, A. Rambaldi, B. Facchinetti, P. Fruscella, O. Spinelli, U. Giussani. *Integrating multiple genetic analysis into the clinical diagnosis of Acute Myeloid Leukemia: a case report.* European Human Genetics Conference, Milan, Italy | June 16-19, 2018.

• M. Tosi, M.L. Guinea Montalvo, **R. Cavagna**, B. Peruta, T. Intermesoli, G. Quaresmini, A. Rambaldi, O. Spinelli. *Digital PCR for Minimal Residual Disease analyses in adult Acute Lymphoblastic Leukemia (ALL).* Foundation FS. XIV Congress of the Italian Society of Experimental Hematology, Rimini, Italy, October 19–21, 2016. Haematologica. 2016;101(s3):S1–S122.

• M.L. Guinea Montalvo, M. Paris, M. Tosi, R.L. Rossi, D. Carpi, B. Peruta, **R. Cavagna**, C. Belotti, P. Zanghì, L. Elidi, S. Salmoiraghi, T. Intermesoli, R. Bassan, A. Rambaldi, O. Spinelli. *Innovative, capture-based Next Generation Sequencing (NGS) approach for Ig/TCR clonal markers identification in adult acute lymphoblastic leukemia patients at diagnosis.* Foundation FS. XIV Congress of the Italian Society of Experimental Hematology, Rimini, Italy, October 19–21, 2016. Haematologica. 2016;101(s3):S1–S122.

• M.L. Guinea Montalvo, M. Paris, M. Tosi, S. Lilliu, B. Peruta, **R. Cavagna**, C. Belotti, P. Zanghì, L. Elidi, S. Salmoiraghi, T. Intermesoli, R. Bassan, A. Rambaldi, O. Spinelli. *Capture-based Next Generation Sequencing (NGS) identifies also rare Ig/TCR rearrangements in adult acute lymphoblastic leukemia patients.* Foundation FS. 21st Congress of the European Hematology Association Copenhagen, Denmark, June 9–12, 2016. Haematologica. 2016;101(s1):1–881.