B-cell malignancies: capture-sequencing strategies for identification of gene rearrangements and translocations into immunoglobulin gene loci

Eileen M Boyle¹,²
Brian A Walker¹
Christopher P Wardell¹
Xavier Leleu²
Faith E Davies¹
Gareth J Morgan¹

¹Centre for Myeloma Research, Institute of Cancer Research, Sutton, Surrey, UK; ²Service des Maladies du Sang, Hôpital Claude Huriez, Centre Hospitalier Universitaire de Lille, Lille, France

Abstract: Aberrant chromosomal translocations involving the immunoglobulin heavy chain (IGH) locus and clonal rearrangements of the variable (V), diversity (D), joining (J) segments (V[D]J) of the immunoglobulin gene are implicated in the initiation of mature B-cell malignancies, including myeloma. Analysis of these events provides information useful for diagnosis and prediction of prognosis, and also provides a useful monitoring strategy for response to treatment in patients with these diseases. Current methods for identification of these events are not generally applicable and give a biased picture of the prognostic significance and clinical relevance of these events. Novel methodologies based on next-generation sequencing are a new and more efficient genetic tool that can be used to screen and characterize these changes at the nucleotide level, offering a deeper and better understanding of the genetic basis of these complex diseases. In this work, we provide a review of the molecular basis of B-cell neoplasms, the methods used to detect them, and how they translate into the clinic.

Keywords: B-cell malignancies, minimal residual disease, molecular, IGH locus, next-generation sequencing

Introduction
Translocations involving the immunoglobulin heavy chain (IGH) locus and clonal rearrangements of the variable (V), diversity (D), joining (J), and constant (C) (V[D]J) segments of the immunoglobulin (Ig) genes are implicated in the initiation of B-cell malignancies, including myeloma. Analysis of these events provides information useful for diagnosing these cancers and for predicting their prognosis after treatment. In this setting, they also provide a tool that can be used to monitor response. Current methods for identification of these events are not generally applicable and give a biased picture of the prognostic significance and clinical relevance of these events. Next-generation sequencing could also be used in a diagnostic setting to identify these changes and may offer additional relevant information for patient management. In this work, we review the molecular basis of B-cell malignancies, the methods used to detect them, and how they might be employed in the clinic.

Immunoglobulin genes and normal B-cell development
The B-cell lineage is derived from lymphoid progenitor cells that differentiate from hematopoietic stem cells. They undergo a complex, highly controlled maturation process leading to expression of a functional Ig formed of two identical heavy chains
(coded by the *IGH* locus) and two identical light chains coded by either of the Ig light chain loci, kappa (*IGK*) or lambda (*IGL*). \(^1\)

The *IGH* locus is located on chromosome 14 at 14q32. This locus comprises approximately 1.5 Mb, and can be divided into four groups depending on the segments of the heavy chain encoded: V, D, J, and C segments. At the pro-B-cell stage, the DNA undergoes sequential DNA recombinations of single D, J, and V segments, which are controlled by the recombination signal sequences that flank each gene segment. This site-specific DNA recombination process is catalyzed by the proteins encoded by recombination activating genes 1 and 2 (*RAG1* and *RAG2*). \(^2\) DNA repair processes such as mismatch repair, base excision repair, and non-homologous end joining, \(^13\) rejoin them. Junction diversity via random addition of nucleotides by terminal deoxynucleotidyl transferase adds further diversity. \(^4\) Once a functional V(D)J rearrangement has been achieved, the other allele is excluded from recombination attempts in a process known as allelic exclusion. \(^5\) The final transcript is composed of the V(D)J rearrangement and the constant region immediately 3′ of the J segment, which in a naïve B-cell is the Cμ region that generates an IgM isotype. \(^6\) Enhancers (E) have been identified. They can be divided into two groups: Eμ enhancers, located in the intronic region between Jμ and Cμ, and a series of enhancers located downstream of Cα in a region referred to as the 3′ *IGH* regulatory region. Eμ is a small 220 bp core element necessary and sufficient to induce transcription, flanked by two matrix attachment regions that contribute positively to Eμ function. They are lineage-specific and are known to drive oncogenesis. \(^7\)

Following the *IGH* locus rearrangement, the light chain loci proceed to rearrange. The kappa and lambda light chain loci are located at 2p12 and 22q11.2, respectively. Like the *IGH* locus, they include V, J, and C segments, but lack D segments. IgK-positive B-cells retain their *IGL* genes in germline configuration whereas the majority of IgL-positive B-cells bear *IGK* rearrangements in line with hierarchical *IGL* chain recombination. First, rearrangements start at the *IGK* locus where Jκ and Vκ segments rearrange with limited junctional diversity. \(^8\) Nonfunctional rearrangements may occur when a downstream element, the kappa-deleting element, rearranges with the Vκ segment or an intron. \(^9\) This leads to inactivation of the kappa locus via deletions of the JκCκ or Cκ regions preventing the expression of that *IGK* allele and therefore participating in the regulation of allelic and κ/λ isotypic exclusion. If the kappa rearrangement is not functional on both alleles, the *IGL* locus proceeds to rearrange. \(^10\)

Finally, pairing of the heavy and light chains results in a functional Ig on the cell surface at the immature B-cell stage, enabling the cell to escape apoptosis and proceed to maturation outside the bone marrow. \(^1\)

After this initial process where the primary Ig repertoire is created, these immature yet immunocompetent B-cells exit the bone marrow and enter the lymphoid organs through the mantle zone into the germinal center (GC), where they proceed to proliferate and undergo affinity maturation after encounter with an antigen. This maturation or GC reaction requires close interaction between B-cells, antigen-presenting dendritic cells, and T-cells. B-cells carrying an antigen-specific B-cell receptor are selected to survive and proliferate. These positively selected cells, or centroblasts, rapidly divide and expand in the dark zone of the GC. Some of them then enter the light zone, stop dividing, and become centrocytes. Centrocytes can revert to centroblasts or terminally differentiate into either memory B-cells or antibody-secreting cells (plasmablasts or plasmocytes). Plasmablasts are short-lived, cycling, antibody-secreting cells that are found in extrafollicular foci. In a poorly understood process, some of them can return to the GC and undergo further clonal selection. \(^11\) The proliferation of naïve lymphocytes occurring during the first encounter with the antigen, the primary immune response, generates IgM or IgD effector B-cells and memory B-cells. The presence of memory B-cells, on subsequent encounters with antigens, enables a quantitative and qualitatively superior response, termed the secondary response. During this GC reaction, two types of DNA modifications occur, class switch recombination and somatic hypermutation. Somatic hypermutation adds point mutations into the variable regions of the Ig genes. \(^12\) As the process of somatic hypermutation is repeated several times, populations of B-cells bearing receptors with increasing affinity toward decreasing levels of antigen are selected. This process is referred to as affinity maturation. Class switch recombination occurs when the V(D)J segment is brought into proximity to another constant region (IGHγ1-4, IGHα1-2, or IGHε) by deletion of the intervening DNA between Cμ and the next constant region. Class switch recombination involves large repetitive switch (S) sequences that are located upstream of each Cμ gene segment. During class switch recombination, the Sμ region recombines with a downstream S region, resulting in a closer proximity of the new Cμ gene segment to the V(D)J region and the *IGH* variable region promoter. \(^13\) Class switch recombination provides an important means of altering the effector function of the antibodies produced by the B-cells and is central to maturation of the antibody response elicited by natural infections and vaccination (Figure 1).
This process occurs during the secondary immune response. The ultimate aim of all these processes is the development of an extensive B-cell repertoire with high affinity antibodies that recognize and optimally bind foreign antigens with effector properties. Although mechanistically different, these two processes are mediated by DNA double-strand breaks induced by transcription-targeted activation-induced cytidine deamination activity. The activation-induced cytidine deamination protein belongs to the APOBEC (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like) family of enzymes.
cytidine deaminases and is capable of deaminating cytidine to uracil in vitro on both single-stranded DNA substrates and single-stranded DNA generated by formation of RNA-DNA hybrids.13,16 Although highly regulated,17 it can induce DNA mutations at a relatively high rate (up to $10^{-4}$ to $10^{-3}$ mutation per base per division) and is a known oncogene.14

Upon terminally differentiating, plasma cells home to the bone marrow where they secrete large quantities of antibodies. Expression of surface Ig is downregulated, hence they no longer respond to antigens. By comparison, the recirculating memory B-cells readily respond to subsequent antigenic challenge as they express functional surface Ig and can, on reactivation, terminally differentiate to antibody-secreting cells.1,11

The process of B-cell terminal differentiation is regulated by the coordinated activity of a small group of master regulatory transcription factors. These factors can be divided into those such as Pax5 and Bcl6 that promote and maintain the B-cell program and others such as Blimp1 (B lymphocyte-induced maturation protein) and IRF4 (interferon regulatory factor 4) that control antibodies secreting cells differentiation. There exists mutually antagonistic transcriptional programs in B-cell and antibody secreting cells that ensure the separation of B-cell and antibody secreting cells fate.18 This process is essential for regulating terminal differentiation.

Immunoglobulin events in B-cell malignancies

Overview

Although essential for an effective adaptive immune response, the different processes that lead to B-cell differentiation are implicated in the pathogenesis of B-cell malignancies. B-cell neoplasms are a heterogeneous group of diseases, each with a distinct level of maturation defined in relation to the GC and their normal cellular counterparts. In broad outline, they can be divided into three types based on their relationship with the germinal center:

- pre-GC center neoplasms with unmutated V segments, eg, acute lymphoblastic leukemia (ALL), mantle cell lymphoma (MCL), and chronic lymphocytic lymphoma (CLL)
- GC type with ongoing V gene mutations, eg, follicular lymphoma and diffuse large B-cell (GC-DLBC) tumors
- post-GC neoplasms with stable mutated V genes, eg, activated B-cell lymphoma (ABC-DLBC), marginal zone lymphoma, and myeloma.19,20

These entities are clonal and are characterized by the presence of a clone-specific rearrangement of one of the Ig gene loci.21,22 V(D)J rearrangements occur early in B-cell ontogeny within B-cell precursors located within the bone marrow23 and can be used to define clonality.22 The second key abnormality at the IGH locus is the presence of reciprocal translocations that deregulate target oncogenes by placing them adjacent to the strong enhancers of the Ig genes.23 These translocations rarely give rise to fusion transcripts. Double-strand breaks and abnormal repair mechanisms contribute to translocations in B-cell malignancies. Depending on the stage they occur at, these translocations are believed to be mediated by different genes, such as the RAG complex at the pre-B-cell and AID at the pro B-cell stage.24 In some cases, these translocations are disease-defining (as in follicular lymphoma, MCL, or Burkitt’s lymphoma),23 in some cases, they have prognostic value such as t(4;14) in myeloma26 (Figure 2).

Immunoglobulin events in pre-GC malignancies

ALL is derived from a normal precursor B-cell. Although long believed to be absent, translocations involving the IGH locus have been implicated in approximately 3% of precursor B-cell ALL. The first recurrent translocation described was t(14;19)(q32;q13).27 Unlike observations in mature B-cell malignancies, the breakpoint on chromosome 19 does not involve BCL3 but CEBPA, a gene of the CCAAT/enhancer binding protein (CEBP) gene family of transcription factors.28,29 Other translocations in ALL involving the IGH locus have been identified, such as t(8;14)(q11;q32), t(14;14)(q11;q32)/inv(14)(q11;q32), t(14;20)(q32;p13), and t(14;19)(q32;q13) with an alternative breakpoint. The partners of such translocations consist of other members of the CEBP family, such as CEBPD, CEBPE, and CEBPG.29 A number of other partners have been described, such as ID4 in t(6;14)(p22;q32)30 and EPOR in t(14;19)(q32;p13).31 Although the partners differ, patients show common clinical features with a pre B-immunophenotype and low blood count. As they remain rare events, their prognostic value is uncertain.32,33 IGH rearrangements are present in up to 90% of ALL blasts. Approximately 50% of them also rearrange the IGK locus and the T-cell receptor (TCR) γδ loci. A combined study of these four loci enables identification of one or more rearrangements in virtually all cases of ALL.34,35 Therefore, IGH rearrangements may be used to monitor minimal residual disease (MRD) and guide treatment strategies, such as allogeneic transplant.

Burkitt’s lymphoma is an uncommon form of non-Hodgkin lymphoma. A defining feature of Burkitt’s lymphoma is activation of the MYC gene located at 8q24 through...
translocation with one of the three Ig loci, which introduces a transcriptional enhancer. In 80% of cases, it involves the IGH locus [(t(8;14)(q24;q32)], in 15% the IGK locus [t(2;8) (p12;q24)], and in 5% the IGL locus [t(8;22)(q24;q11.2)]. The breakpoint occurs between the first (non-coding exon) and second exon (where the start codon is located), leading to an enhancer-driven activation via a cryptic promoter (P3). IGH rearrangements are also present in virtually all patients.36

The t(11;14)(q13;q32) translocation that juxtaposes the proto-oncogene CCND1 at 11q13 to the IGH locus at chromosome 14q32 is considered the primary oncogenic event in the development of MCL. It occurs at the pre-B cell stage of differentiation during recombination of the V(D)J segments of the IGH variable region.37 However, the tumor is composed of a specific population of mature B lymphocytes, indicating that the full neoplastic phenotype is acquired at a later stage of differentiation. Historically, naïve B-cells have been considered the normal counterpart to MCL tumor cells based on IgM/IgD and CD5 expression, their location in the mantle zone, and early descriptions of the predominantly unmutated variable region. More recently, comprehensive analysis of the B-cell receptor diversity in MCL has shifted this view to a more complex ontogeny, in which antigen selection plays a role. Recent studies have also shown that the subset of cases that carry somatic hypermutation (15%–40%) also have an indolent phenotype and clinical behavior.38–40

Rearrangements of the IGH locus with oncogenes appear to be rare in CLL. Rare cases of t(14;18)(q32;q21) involving the IGK locus and BCL2, have been described. The t(14;19) (q32;q13) involving the BCL3 gene was found in six of 4,487 cytogenetically analyzed lymphoproliferative disorders. Five of these six cases were eventually classified as CLL.41 Initially, it was thought that the clonal accumulation of CD5+ B-cells in CLL was comprised of naïve B-cells with unmutated IgH variable region. However, recent studies showed the presence of significant somatic mutations in approximately 50% of patients, suggesting the neoplastic cells may correspond to post GC memory B-cells.42–44

Immunoglobulin events in GC malignancies

The hallmark of follicular lymphoma, t(14;18)(q32;q21) results in a constitutive overexpression of BCL2, allowing B-cells to abrogate the default GC apoptotic program. Nonetheless, this translocation is neither necessary nor sufficient for diagnosis, as it is absent in 15% of follicular lymphomas and present in about 20%–30% of GC B-type
DLBCL lymphomas. The somatic hypermutation process is ongoing in follicular lymphoma, and recent studies suggest that a disruption in this process occurring late in evolution of the disease may select a more aggressive disease and lead to transformation into high-grade lymphoma.

GC B-type DLBCL lymphoma seems to originate from the GC and frequently shows ongoing somatic hypermutation. Forty-five percent of patients harbor a t(14;18)(q32;q21) leading to a constitutive activation of BCL2, a feature that is absent in ABC-like disease.

**Immunoglobulin events in post-GC malignancies**

ABC-DLBCL lymphoma seems to originate from a B-cell that is in the process of differentiating into plasma cells. Translocations have been described, but unlike GC B-type DLBCL lymphoma, they do not harbor any t(14;18)(q32;q21).

The chromosomal translocations t(1;14)(p22;q32), t(14;18)(q32;q21), and t(3;14)(p13;q32) are all known to occur with variable frequencies in mucosa-associated lymphoid tissue (MALT) lymphomas, resulting in IGH-BCL10, IGH-MALT1, and IGH-FOXP1 rearrangements, respectively. The t(1;14)(p22;q32) and t(14;18)(q32;q21) occur in a small subgroup of almost exclusively non-gastric MALT lymphomas.

Myeloma is a late B-cell malignancy where malignant plasma cells expand and accumulate in the bone marrow. As IgM and IgD are rare, myeloma is considered a post-class switch disease. Translocations involving the IGH locus occur in approximately 50% of cases and include t(4;14)(p16;q32) (15%), t(11;14)(q13;q32) (20%), t(14;16)(q32;q23) (<5%), t(6;14)(p21;q32) (<5%), t(8;14)(8q24;q32) (<5%), and t(14;20)(q32;q12) (<5%). They are responsible for mediating ectopic expression of specific genes brought under the control of strong enhancers. The t(11;14)(q13;q32) leads to overexpression of CCND1, t(6;14)(p21;q32) deregulates CCND3, and t(4;14)(p16;q32) involves WHSC1 (also known as MMSET) and FGFR3. The majority of the breakpoints are localized within the switch region and are generated through class switch recombination. All t(4;14) are mediated through class switch recombination; however, 21% of t(11;14) and 25% of t(14;20) are generated through DH-JH recombination, with breakpoints within the D4-JH region. These two groups also generate translocations through receptor revision, as determined by the breakpoints in 10% and 50% of t(11;14) and t(14;20) samples, respectively. These data suggest that, although similar, they would involve different genes. The RAG complex would be responsible for the translocations occurring through DJ rearrangements via the pro-B-cell stage in the bone marrow or through an aberrant RAG expression in the post-GC. AID would be responsible for the class switch recombination-mediated translocations. In MCL, although the translocation partners are similar, all the t(11;14) are generated through V(D)J rearrangement, suggesting the role of RAG at the pro-B-cell stage. Deregulation of MYC by t(8;14)(q24;q32), t(2;8)(p12;q24), and t(8;22)(q24; q11.2) is also present at diagnosis. Similarly, MYC translocations may also involve non-Ig partners. These translocations function in the same way by placing an enhancer from a B cell lineage gene next to MYC, leading to its overexpression. Comparable mechanisms are seen in other B-cell malignancies.

IGH rearrangements are detectable in most tumor cells in a patient, suggesting they are initiating events. Most plasma cells secrete an Ig coded by a functional hypermutated V(D)J. The rearrangement is often biallelic and the other allele can carry either a complete but non-functional rearrangement (40%) or a single DJ rearrangement in the same frequencies as those observed in normal B-cells, supporting the principle of allelic exclusion. Hypermutation of the rearranged V(D)J is present in most cases. The level of hypermutation ranges from 9% to 23%, which is consistently higher than in other diseases such as CLL, hindering a polymerase chain reaction (PCR)-based molecular diagnosis approach. Of note, somatic hypermutations are absent on the non-completely rearranged allele. This allele may therefore be used for PCR. As seen in other B-cell malignancies, the IGK locus is rearranged in kappa myeloma and IGL in the lambda isotype myeloma.

**Clinical value of detecting clonality and translocations**

**Detection of clonality**

The concept of clonality is based on the hallmark of cancer where, once a malignant B-cell has rearranged its Ig genes, every daughter cell will bear the same rearrangement. IGH rearrangements are not B-cell-specific, as they can occur in 22% of T-cell ALL, 8%–16% of T-cell lymphomas, and 4%–18% of AML. Nevertheless, they are crucial in B-cell malignancies because their identification can be used to confirm the diagnosis and guide follow-up. Despite the well established morphological and phenotypic diagnostic criteria, definite diagnosis of a suspected lymphoproliferative disease often requires clonality testing. Retrospective studies suggest that 5%–15% of suspected lymphoproliferative disease diagnoses cannot be made based on morphology or immunophenotyping alone. This percentage
increases in immunocompromised patients.\textsuperscript{57,59} The ability of demonstrating V(D)J rearrangement is an effective strategy to demonstrate clonality;\textsuperscript{59} nevertheless it varies among B-cell neoplasms.\textsuperscript{57,58,60,61} V(D)J rearrangements may also be used to establish the clonal relationship between two coexistent lymphoid malignancies.\textsuperscript{62–64}

Beyond diagnosis, these rearrangements may also be used to monitor MRD. Monitoring MRD is a key requirement for the management of most hematological malignancies as it can be used:

- as an early surrogate marker of treatment efficacy that can therefore be used to guide treatment strategies early on, such as de-escalation strategies leading the way into personalized medicine
- as a prognostic marker that can predict outcome
- as a follow-up marker for remission control that can be used as a trigger for pre-emptive treatment.

Evaluation of MRD by multiparameter flow cytometry requires that neoplastic cells express an antigen profile that differs from that observed in normal hematopoietic cells and that this pattern of expression is not affected by chemotherapy. This immunophenotyping approach was to some extent successful because it was able to predict outcome in diseases such as pediatric ALL,\textsuperscript{65} myeloma,\textsuperscript{66} and CLL.\textsuperscript{57} It was also used to guide treatment approaches, notably in pediatric ALL where MRD could be used to determine requirements for an allogeneic transplant\textsuperscript{65} or CLL where it can be used in a risk assessment strategy to guide treatment after allogeneic transplant.\textsuperscript{68} With the introduction of standardized quantitative PCR for the measurement of abnormal rearrangements in ALL,\textsuperscript{69–72} it appeared that the extent and depth of response based on a molecular approach was even more prognostically informative than that observed by flow cytometry.\textsuperscript{67,68} With increased sensitivities in comparison with immunophenotyping, in ALL, molecular MRD can distinguish prognostic subgroups.\textsuperscript{73–77} Although rare cases have a BCR-ABL or a MLL-AFA4 fusion transcript, for technical convenience, the rearrangements more commonly studied for follow-up of MRD are those of TCRγδ, IGH, and IGK. MRD technology has more recently been used in the study of Burkitt’s lymphoma\textsuperscript{78} and MCL.\textsuperscript{79} In myeloma, there is currently no standard molecular MRD detection as PCR-based assays for V(D)J rearrangements are hindered by somatic hypermutation, given the non-annaling of the PCR primers.\textsuperscript{54}

Detection of somatic hypermutation

This is achieved by sequencing the IGH variable region segment of a clonal population and comparing it to a database of germline sequences. If the difference accounts for more than 2% of the nucleotides, the segment is considered mutated. In fundamental studies, Damle et al and Hamblin et al have shown\textsuperscript{50,61} that the presence of unmutated IGH variable region genes predicts for an inferior survival in CLL. There is to the extent of our knowledge no equivalent study in myeloma.

Detection of translocation

Translocations in B-cell malignancies can be used to confirm the diagnosis, predict outcome, and maybe monitor MRD. The clinical value of detecting a chromosomal translocation differs according to disease type. They are required for diagnosis in diseases such as MCL [t(11;14)(q13;q32)], follicular lymphoma [t(14;18)(q32;q21)], and Burkitt’s lymphoma [t(8;14)(q24;q32)].

In a clinical setting, apart from diagnosis, the main reason for identifying translocations is to stratify patients and use this information to guide treatment. In myeloma, for instance, some translocations are of prognostic value, such as t(4;14)(p16;q32), and to a lesser extent t(14;16)(q32;q23) or t(14;20)(q32;q12). When used in conjunction with other biologically relevant elements such as beta-2 microglobulin, they can be used to define molecular subgroups with similar outcome that would benefit from specific treatment options.\textsuperscript{26} Finally, preliminary data suggest that detection of translocations or the fusion transcript that may arise from these balanced translocations can be used to detect MRD in Burkitt’s lymphoma.\textsuperscript{78}

Traditional methods for detecting immunoglobulin loci rearrangement events

Identification of V(D)J rearrangement

V(D)J rearrangements were first assessed by Southern blotting\textsuperscript{21} with hybridization probes mapping to the V, J, or C segment.\textsuperscript{59} This is traditionally regarded as a time-consuming, labor-intensive, and technically demanding procedure. It also requires both high quantities and good quality of DNA, so cannot be performed on fixed samples. False positives are rare, and false negatives account for 5% of samples. As it has a lower sensitivity than PCR, this method has gradually been replaced by PCR.

The PCR approach relies on amplification of the IGH locus using conserved sequences. It is believed to be both more rapid and less labor-intensive than Southern blotting. It also requires lower qualities of DNA and can be performed on previously fixed or embedded samples. It is based on the IGH
variable region and JH consensus primers FR I, II, III, and IV. Using a single FR yields sensitivity of 50%–75%. When combining multiple primers or adding the IGK primers, sensitivities are greater than 90%. BIO MED-II primers target FR I, II, and III, the incomplete DJ rearrangement, and the IGK locus. Adding the DJ rearrangement primers targets the non-coding allele. This is of particular importance in GC and post-GC malignancies. As it does not undergo somatic hypermutation, primers have a higher probability of binding appropriately. The IGK primers bind the VJ segments as well as the kappa-deleting element. Adjunction of IGL primers does not seem to improve sensitivity. False positives will arise in the event of oligo or pseudoclonal populations. This is observed when either the sample or the repertoire is poor (eg, immune-compromised patients or post-treatment samples). False negatives relate to ineffective binding and subsequent lack of amplification. This is chiefly observed in three different circumstances: in the presence of somatic hypermutation (in GC or post-GC malignancies), in the case of a translocation breakpoint involving the consensus region, and in the case of small degraded DNA samples (<200 bp). PCR is able to yield results in up to 90% of cases of CLL and 98% of ALL cases. Nevertheless, in post-GC malignancies that have undergone somatic hypermutation, such as DLBC lymphoma, follicular lymphoma, and multiple myeloma, the rates of failure are higher (15% in follicular lymphoma and up to 60% in myeloma).

**Identification of translocations**

Karyotyping and conventional cytogenetics made their way into the clinic with the development of more effective techniques to culture bone marrow and blood and improved metaphase preparations. More accurate identification of individual chromosomes became possible with banding techniques, first using fluorochromes (Q-banding) and then enzymatic treatment strategies (G-banding). Although standardization facilitated the description of chromosomal abnormalities, these techniques lacked sensitivity and were not appropriate for tumors with low proliferating indices in vitro. A conventional cytogenetics approach is currently standard for the detection of established chromosomal abnormalities in ALL within a routine clinical environment. Nevertheless, it fails to identify morphologically cryptic or complex translocations. It also has a high failure rate inherent to this technique (requires live cells) and to the tumor characteristic (high failure rates in tumors with low proliferation indices). The latter explains the limitation of such an approach in myeloma. The performances of karyotyping were improved by addition of metaphase fluorescence in situ hybridization (FISH) that helped to identify specific fusions without resolving the technical limitations.

FISH is based on annealing of a DNA probe to a complementary sequence on a target genomic DNA. The use of split signal probes has led to the identification of morphologically cryptic translocations. Interphase FISH on selected cells increased the sensitivity of such approaches. Although sensitivity is increased, this method remains labor-intensive with high failure rates and is inherently biased toward a small selection of translocations.

Whole genome DNA analysis using comparative genomic hybridization was another significant technical development in molecular cytogenetics. It can detect abnormalities arising from the gain and loss of DNA. High resolution array comparative genomic hybridization may detect submicroscopic changes in DNA copy number to a resolution <1 Mb. This approach enables us to accurately locate the breakpoints involved in such rearrangements. A parallel technology utilizing single nucleotide polymorphism arrays has been used to show copy number changes and loss of heterozygosity. These arrays can also detect translocations if they are unbalanced.

There is a need for a high throughput tool to detect these events at the molecular levels, which could be used to better understand the mechanisms of disease, its heterogeneity, and potential new targets.

**High throughput strategies**

Here we provide a description of the capture-sequencing method, its role in better characterizing IGH alterations, and its potential applications and limitations in the study of mature B-cell neoplasms.

**Principles of next-generation sequencing**

The main principle of next-generation sequencing is based on massively parallel sequencing where small DNA fragments are sequenced in parallel (Figure 3). Bioinformatics is used to piece together these fragments by mapping these fragments or reads to the reference genome. Each base of the genome is sequenced multiple times, giving additional depth that translates into accuracy of the data thus generated. This high throughput method can be used on entire genomes, individual genes, or the 25,000 coding genes (whole exome). This method can provide insight into a wide panel of DNA variations, such as missense mutations, small insertions or deletions, and structural changes, such as inversions, translocations, and rearrangements. Copy number changes and
Capture-sequencing in B-cell malignancies

microdeletions can also be derived from the data. It does not require any pre-existing knowledge of the gene or its location, and thus offers an unbiased view of the whole exome. The main limits of this approach are the relative error that occurs in the event of a high GC content or a repeat architecture, the infrastructure required to comprehensively analyze the sample, and the cost that limits its implementation in the clinic. Targeted capture approaches based on RNA baits mapping to a specific region offer a methodology to enrich and select specific regions of interest. This has been used in the research setting to define the translocation breakpoints in myeloma and in DLBC lymphoma. As the technique requires small amounts of DNA (50 ng) and can be performed on formalin-fixed, paraffin-embedded extracted DNA, it provides an opportunity for rapid detection of translocation in the clinical setting.

Clonality detection using immunoglobulin gene rearrangements

By defining a capture for the Ig gene loci, high-throughput sequencing methods can be applied to universally amplify these segments and identify all clonal gene rearrangements despite somatic hypermutation. It is also possible to amplify the Ig rearrangements using PCR followed by massively parallel sequencing of the product to detect clonality in B-cells. This provides a more robust and more specific means of detecting subclones emerging through disease progression. Measurements may be repeated at different time points, such as at diagnosis or disease progression. This approach has been tested in ALL where it was able to detect MRD. As numerous correlative studies suggest, in ALL the prognostic power of MRD could also increase due to the higher sensitivity of this approach. Studies in CLL suggest it could be a good diagnostic tool and useful for monitoring MRD after transplant. In myeloma, 133 newly diagnosed patients with a very good partial response were evaluated using a next-generation sequencing-based MRD test. This approach proved feasible (91% of patients were evaluable) and the median overall survival was proportionate to the depth of response. Studies in other B-cell malignancies are encouraging, but should nevertheless be confirmed in a prospective setting. The current sequencing costs for rearranged antigen-receptor genes are comparable with those of flow cytometry and PCR. Sequencing data may be complex to interpret, but the current analytic algorithms facilitate detection and quantitation of genetic events.

Detecting chromosomal translocations into immunoglobulin gene loci

The identification of IGH translocations requires firstly the design and fabrication of a targeted capture library directed at the IGH locus. Although not currently available on a routine basis, this technique could be used to detect translocations by detecting the presence of reads mapping to different chromosomes. Exact breakpoints could be identified and could be used to determine PCR-based assays for MRD.

Discussion

Next-generation sequencing-based tests could be implemented to replace the current PCR-based assays for
MRD detection. Next-generation sequencing methods yield higher sensitivities in the MRD setting than PCR in most cases of B-cell neoplasms in which they have been tested. When tested in 106 ALL patients and compared with flow cytometry, next-generation sequencing was able to identify all positive patients and an additional ten patients previously considered negative. In the same cohort, when compared with PCR, it was able to identify 97% of positive cases and an additional three patients.93 Similar results were seen in CLL and other B-cell malignancies, thus adding sensitivity and precision to MRD detection. Nevertheless, these data need confirmation in prospective randomized trials in order to determine the exact prognostic significance of next-generation sequencing and how it should be used to guide treatment.91

The capture sequence strategy may also be used to identify translocations. Whole exon sequencing has been used both in myeloma and DLBC lymphoma. New translocation partner chromosomes and exact breakpoints were identified. Furthermore, by providing information on both V(D)J rearrangement and translocations, capture sequencing strategies may also offer additional information by not only determining the breakpoint but also the putative mechanism of the chromosomal event (aberrant V(D)J rearrangement or class switch recombination, for instance). From there it was shown that, in some cases, the initiating events in post-GC-type disease, such as myeloma, occurred earlier in B-cell development (at the pre-GC stage).32 This may be of clinical relevance.

Smaller targeted panels may be used in a diagnostic setting. As the capture sequence strategy is based on DNA rather than viable cells or easily degraded RNA, it would be more reliable and easier to implement, especially in the case of paraffin embedded samples.

Finally, next-generation sequencing may be used to monitor clonal evolution. Firstly, the study of somatic hypermutation and its progression throughout disease could offer further insight into the clonal architecture and evolution of B-cell malignancies, especially those with ongoing somatic mutations, such as follicular lymphoma or myeloma. Secondly, the identification of multiple subclones and their evolution throughout treatment and follow-up may be of additional prognostic value.

Conclusion

The development of molecular diagnostics has significantly improved the detection of IgE events such as translocations and V(D)J rearrangements in B-cell malignancies. From a clinical perspective, this has translated into more accurate diagnosis, better risk stratification, and more precise monitoring of patients. These studies have also given us a better understanding of the ontogeny of B-cell malignancies. The translation of the more novel approaches into the clinic, such as next-generation sequencing, could also benefit patients. Not only are they more sensitive than standard approaches in the detection of these events, but they could also give additional information on the clonal architecture of these malignancies that could translate into the clinic.

Disclosure

The authors have no conflicts of interest in this work.

References
