Diagnostic molecular techniques in haematology: recent advances

Aikaterini Koutsi, Elisavet-Christina Vervesou

Department of Hematology, Errikos Dunant Hospital Center, Athens, Greece

Contributions: (I) Conception and design: A Koutsi; (II) Administrative support: A Koutsi; (III) Provision of study materials or patients: A Koutsi; (IV) Collection and assembly of data: A Koutsi; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Aikaterini Koutsi, MD, BS. Department of Hematology, Errikos Dunant Hospital Center, 107 Mesogeion Avenue 115 26, Athens, Greece. Email: koutsikatrina@gmail.com.

Abstract: Hematopoietic disorders are often driven by genetic mutations and epigenetic alterations. New advanced technologies including next-generation sequencing, ultra-deep PCR and whole-genome and exome sequencing were proved very efficient in detecting several mutations implicated in the pathogenesis of hematological diseases. Emerging evidence indicates that genomic data can be useful in all aspects of clinical practice including diagnosis, prognosis and prediction of response to specific treatments, as well as in the development of novel targeted treatments for patients with hematological disorders.

Keywords: Hematology; sequencing; target therapy; mutations; diagnosis

Submitted Mar 20, 2018. Accepted for publication May 17, 2018.

doi: 10.21037/atm.2018.05.30

View this article at: http://dx.doi.org/10.21037/atm.2018.05.30

Introduction

Benign and malignant hematological disorders are heterogeneous in both biological and clinical aspects. The alterations of genomic profile associated with these diseases are complex and variable including mutations, translocations, karyotypic rearrangements and post-translational modifications. In some cases, several genetic changes are required, to induce the onset of disease. This evidence in association with the evolution of molecular techniques has led to a modification of the existing dogma focusing on a single gene or single pathway analysis (1). The development of new methods in molecular biology has not only allowed the individualized molecular diagnosis of diseases but has also led to the discovery of genetic or targeted therapeutic schemes with cytotoxic, anti-metabolic or immunomodulatory properties.

Hematological physiology and pathology, independent of whether it is aggressive or indolent, affect patients of all ages with numerous clinical presentations. The concept of this review is to analyze the molecular basis of hematological diseases, as well as to present some new molecular technology and how they can affect overall survival. Using polymerase chain reaction (PCR), karyotype analysis, fluorescence in situ hybridization (FISH) and next-generation sequencing (NGS) it is possible to design better risk stratification categories and determine minimal residual disease (MRD). Immune check points inhibitors, antibodies and chimeric antigen receptor (CAR)-T cells can guide most efficient therapeutic strategies.

Understanding the molecular basis of hematological diseases

Normal cell life is highly dependent on gene expression and any qualitative or quantitative alterations on the cascade of genetic information as well as changes in the time frame of gene activation, lead to inappropriate protein production (2).

Such changes induce irregular survival abilities, inappropriate response to external signals, autonomous amplification and deregulation of apoptosis pathway, formation of autocrine loops and promotion of angiogenic pathways. However, despite those irregularities, the mutated cell has a selective advantage.
Oncogenes

Oncogenes refer to mutated genes generated from proto-oncogenes coding for proteins that regulate proliferation and differentiation or enhance epigenetic modifications.

They are usually growth factors or mitogens secreted by cells with autocrine or paracrine properties. Such an example presents the c-Sis oncogene. Oncogenes can be receptor tyrosine kinases including epidermal growth factor receptor (EGFR), platelet-derived growth factor receptor (PDGFR), vascular endothelial growth factor receptor (VEGFR) and human epidermal growth factor receptor 2 (HER2/neu) (3,4).

Oncogenes may have the role of cytoplasmic tyrosine kinases. In this category, belongs the Abl gene in chronic myeloid leukemia (CML) (the Philadelphia chromosome) and the Src family, Syk-ZAP-70 family and, Bruton's tyrosine kinase (BTK) family of tyrosine kinases. All these proteins constitute significant targets for treatment. The identification of Bcl-Abl fusion gene, was the first chromosomal abnormality associated with a specified disease and with the arrival of the Abl tyrosine kinase inhibitors (TKIs) in the history of CML turn over (5,6).

In the TKI era, the overall survival probability of patients with CML was raised to 92% (7). On the other hand, BTK is a tyrosine kinase involved in the transmission of different intracellular signals that reflect B cell physiology. The B cell receptor (BCR) can activate BTK and present an attractive therapeutic target for B cell disorders.

The BTK inhibitors are used in B-cell malignancies, especially in patients with chronic lymphocytic leukemia (CLL), mantle cell lymphoma (MCL) and Waldenstrom's macroglobulinemia (WM) (8). The role of oncogene does not end here, the Raf kinase and cyclin-dependent kinases are cytoplasmic serine/threonine kinases, the Ras protein is a regulatory GTPase (9,10). Ras influences major signaling pathways that lead to cellular growth and proliferation. Finally, an oncogene may be a transcription factor, such as the myc gene, which regulates the transcription of genes that induce cellular proliferation. C-Myc is a key therapeutic target in high-grade (“double hit”) diffuse large B-cell lymphomas and T-cell lymphomas (11,12). Double-hit lymphoma (DHL) is a B-cell lymphoma with MYC/8q24 rearrangement plus BCL2 and/or BCL6 rearrangements using cytogenetic studies. The detection of those chromosomal rearrangements expressed concomitant influence therapeutic strategies and reflect patient survival (13).

Tumor-suppressors genes

Tumor-suppressor genes encode for proteins that participate in the cell cycle. They can be receptors for different growth factors or may play the role of enzymes that control DNA repair. Loss of expression of those genes is associated with high risk of developing a malignancy. The first tumor suppressor gene was identified by studies on retinoblastoma (RB). The function of Rb as a tumor suppressor gene was validated by studies investigating the loss of normal Rb allele. Isolation of the Rb gene, as a molecular clone in 1986, demonstrated that Rb is lost or mutated in RBs.

Gene transfer experiments clarify that introduction of a normal Rb gene into RB cells cancels their tumorigenicity, indicating the activity of Rb as a tumor suppressor (14,15). Nodal role, in myeloproliferative and lymphoproliferative disorders, is played by the mutations in p53 protein, a nuclear transcription factor with a pro-apoptotic function, able to interrupt the cell cycle in G1 in response to damaged DNA and required for apoptosis induced by a variety of stimuli.

The mutations of p53 result in loss of function and are restricted within the DNA-binding domain of p53. p53 is frequently inactivated in different human malignancies, including leukemia, lymphomas, sarcomas, brain tumors, and carcinomas of several tissues, including breast, colon, and lung. In total, mutations of p53 may be involved in up to 50% of all cancers, making it a universal target of genetic alterations in malignancies (16,17).

Epigenetics alterations

Most important epigenetics alterations are DNA methylation, histones modifications (acetylation or methylation) and microRNA (miRNA) regulations (18). They determine gene transcription via different mechanisms. The methylation of CpG rich part of a promoter influence the downstream coding region. Degree of DNA acetylation reflect chromatin condensation. MicroRNA (miRNAs) are non-coding RNAs that manage gene expression via mRNA degradation or translational repression (19). Epigenetics changes are identified in leukemia and myelodysplastic syndromes. Those alterations are an attractive target for treatment as are the inhibitor of DNA methyltransferase (DNMTi) and the inhibitor of histone deacetylase (HDACi) (20).
Molecular methods

The basic principle of molecular methods involves isolation, designation and manipulation of genes. All methods involve the extraction, isolation of nucleic acids and further separation of DNA, RNA and proteins using ribonuclease, proteolytic enzymes and different detergents.

PCR and prospects

The PCR first described in 1983 remains the most popular in vitro technique for take advantage of all the information that DNA can offer in diagnosis or treatment. It is based on the ability of DNA polymerase to synthesize a new strand of DNA, complementary to the offered template strand. PCR advantages include high specificity and fidelity while its limitations include depletion of reagents, the time scale of running programs and interpretation of false positive results due to contaminated materials.

PCR products can be sequenced directly or used in recombinant DNA technology. Once PCR was discovered and different parameters were studied, a slew of variation appeared. At the same time, software evolution and data processing guarantee the incorporation of additional information. It is therefore possible to design a whole chromosome with overlapping primers and screen it for the presence of deletion mutations.

Another application is the anchored PCR, which uses a primer of known sequence to investigate and analyze DNA of unknown sequences. Additionally, in the bubble PCR, one primer is like a bubble and can detect a fusion gene.

In inverted PCR, two different steps exist, the circulation of the fragments accompanied by their split at different sites to amplify the unknown region between two primers. Finally, the multiplex PCR can efficiently amplify several fragments simultaneously. The simple sequence length polymorphisms (SSLPs) presents a method that can be applied when the sequence designed holds an internal tandem repetitive sequence and the PCR products are polymorphic. This internal tandem can be a genetic marker. Another significant evolution was the reverse transcription-PCR (RT-PCR), which is restricted on the expressed genes. The RT-PCR is the most sensitive technique for mRNA detection. In fact, this technique is quite specific to allow quantitation of RNA from a single cell. The only limitation is the short stability of RNA. Quantitative PCR (qPCR) and RT-qPCR can be applied to detect, characterize and quantify nucleic acids. However, in qPCR, fluorescent labeling makes possible the data collection during PCR reaction with fluorescence being calculated during each step. The fluorescence signal is amplified proportionally to the replicated DNA, and therefore DNA concentration is evaluated in real time. In probe-based qPCR, many targets can be simultaneously recognized in each sample. Assays can combine high sensitivity with reliable specificity. Real-time PCR detects and analyses the result as a ratio to a standard housekeeping RNA, such as ribosomal RNA or glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA. The ratio can be further converted to the number of mRNA molecules present in the test sample (21-26).

The recombinant DNA techniques can be applied in clinical medicine, genetics and forensics. Myeloproliferative neoplasms are disorders associated with single nucleotide polymorphism (SNP), for example, the mutation in PDGFR alpha (PDGFRA) in eosinophilia and the Janus kinase 2 (JAK2) and myeloproliferative leukemia protein (MPL) in essential thrombocythemia. The identification of one mutation help to distinguish the disorder. Concerning another category of hematological malignancies, the myelodysplastic neoplasms, the genetic profile has an impact in treatment response and risk of transformation in acute myeloid leukemia (AML).

Cytogenetic techniques

Karyotypes expose characteristic structure configurations for each chromosome. Clinical cytogeneticists investigate human karyotypes to uncover genetic changes in chromosome number deletions, duplications, translocations or inversions.

Certain chromosomal translocations are considered pathognomonic of specific diseases and others are used to risk stratifications of different disorders as lymphomas, myelodysplastic syndrome, multiple myeloma (MM) and AML. The karyotype is analyzed at the metaphase stage of cell division. Limitations of this method include the necessity to have active mitotic cells while abnormalities of few bases cannot be detected (27-29).

FISH

FISH relies on the complementary nature of DNA or DNA/RNA double strands. Selected DNA strands combined with fluorophore-coupled nucleotides count as probes to hybridize onto the complementary sequences in tested cells and tissues and can be further visualized at a
fluorescence microscope or an imaging system.

FISH tests using panels of gene-specific probes for deletions, amplification, and translocations are used for the study of hematologic and solid tumors, detection of intracellular microorganisms and parasites. The advantage or this method is the capacity to identify one sequence without disrupting the nuclear genome and the possibility to measure mRNA expression of multiple genes within one single cell.

FISH may yield false negative results when genomic imbalances are smaller than the size of a FISH probe or chromosomal rearrangements are complex. Furthermore, false positive results can be retrieved when two fluorescent signals co-localize due to viewing of a three-dimensional (3D) nucleus in two dimensions (30,31). Microarray technology is exponentially developed following whole genome sequencing.

Based on the rules of hybridization, specific sequences are immobilized to a surface, reacting with labeled cDNA targets. A signal is produced post hybridization. The array can be applied to quantify many different nucleic acid sequences in the same time. The methods of high-resolution genome arrays, including comparative genomic hybridization (CGH) and SNP, guarantee the detection of short number variations of clinical significance, in addition to most abnormalities identified by G-banding and FISH.

SNP-based genotype at a submegabase resolution, promotes the detection of small areas of genomic loss of heterozygosity (LOH) or uniparental disomy (UPD). The data collected via those methods can be analyzed by different software. The genotypes at each site can be used in genome wide association studies (GWAS), in which the allele frequencies at each SNP are compared in disease cases and unaffected controls. Another interesting point is that DNA changes remain proportional to the intensity of fluorescent signals (32-35).

The list of the chromosomal translocations associated with a hematological disorder is without end. The remission, relapse and survival are measured by the presence of cytogenetic abnormalities. Applying FISH method during the follow-up of patient with AML, acute lymphoblastic leukemia (ALL), CLL and MM MRS can be detected, and another line treatment can be decided. Interphase FISH offer the opportunity to assure the follow-up in disorders with a low mitotic yield as for example disorders included in B cell non-Hodgkin lymphoma (NHL).

The clinical hematology guides lines are changed since patient can be stratified in favorable intermediate or unfavorable categories according to cytogenetic and molecular findings.

NGS

The length of fragment and the long time required for Sanger sequencing, were the firelight to search for a sequencing technology capable to produce large amount of data in a fast and low-cost manner. NGS can perform sequencing of millions of small fragments of DNA in parallel. There exist different platforms including Illumina (Solexa), Roche 454, Ion torrent: Proton/PGM and SOLiD sequencing.

In NGS, many short fragments of DNA are sequenced in parallel. The first step is to cleave the DNA, each fragment becoming the template for a PCR. For sequencing, optical signals or H+ ion occur from addition of a dNTP to a DNA polymer. NGS is applied for sequence format, alignment of Next-Gen sequences to reference sequences, identification of variants, such as SNPs and structural variations and sequence assembly.

The data obtained from NGS can provide a comprehensive assessment of the genomic landscape correlated with the genesis and evolution of different malignancies. Molecular diagnostic assays for hematological malignancies have been focused typically on the discovery of mutations of pre-identified genes (36-38).

Proteomics

Cell phenotype is highly dependent on proteins. Many proteins undergo extensive posttranslational modifications that influence their activity and function, including cleavage, chemical modifications such as phosphorylation and glycosylation, and interaction with other proteins. The “proteomics” came to fight genome analyses limitations. The term “proteomics” is introduced in 1995 and was determined as the large-scale characterization of the entire protein complement of a cell line, tissue, or organism.

The main role of proteomics is to acquire a global, integrated view of biology by studying all the proteins transduced in a cell and generate a 3D map. Information can be taken for protein-protein interactions and protein modifications induced from intracellular or extracellular signals affecting their function. This analysis is based on mass spectroscopy according to a complex bioinformatic program. This technique requires conversion of samples to ions.

Two methods are commonly used, electrospray ionization (ESI) and matrix-assisted laser desorption/ionization
(MALDI) that guarantee the integrity of peptides. “New generation” proteomics are the protein microarrays that can be applied for detection of protein-protein, protein-lipid, protein-nucleic acid, and enzyme-substrate interactions (39,40).

**Clinical applications of molecular technology in hematology**

Molecular biology techniques are used to study molecular pathways of the central dogma of life, the replication, transcription and translation of genetic material. Life requires storage and replication of genetic information as well as the ability to transmit information. The manipulation of DNA drive to a non-limit application, displays an important impact in medicine and pharmacology, from diagnosis of diseases to manipulation of the genome, generation of specific antibodies and targeted treatment.

**Inherited mutations**

**Hemoglobinopathies and thalassemias**

Hemoglobinopathies and thalassemias are well-studied diseases by molecular techniques. Apart from the common variants, using allele-specific primers, uncommon mutations (Hb Q-India, HbNedlands, Hb Queens Park) are identified, allowing the possibility to proceed with a family or a prenatal screening study.

Applying gap-PCR with primers that can bind to both deletions sides, α-thalassemias can be successfully diagnosed, usually revealing one or more variably sized deletions of α-globin gene (41,42).

**Pharmacogenomics**

Pharmacogenomics present a part of pharmacology dedicated to the description of genome variants that reflect drug behavior, typically via alterations in drug’s pharmacokinetics (absorption, distribution, elimination, metabolism) or via accentuation of drug’s pharmacodynamics (modifying drug’s target or perplexing biological pathways that remodel sensitivity of drug’s pharmacological effects).

For hematology, significant variations include the deficiency of glucose-6-phosphate dehydrogenase (G6PD), the critical position of the family of isozymes cytochrome P450 and the impact of HLA-B allele. The genetic polymorphism in thiopurine methyltransferase (TPMT) detected by PCR methods results in loss of function, predisposing all patients to severe hematologic toxicity unless the dose of mercaptopurine is reduced.

Using a range of techniques, it is possible to identify mutations more or less common in one population (43-46). Recognizing whether a genomic variant is translated in a mutated enzyme substrate and acts like an inducer or inhibitor can prevent unexpected toxicity, serious adverse events and significant interactions.

It can also avoid co-administration and adjusting of a patient’s drug dose to guarantee an optimal response with minimal adverse effects. It is important to mention that the hematological malignancies are designated by genes alterations or single nucleotide mutations, which can influence drug’s pharmacokinetics and/or pharmacodynamics.

Known examples are the Bcr-Abl fusion gene in CML, its mutation and selection of the best TKI, the FLT-ITD mutation or PML-RARA fusion gene in AML, the activity of sorafenib or retinoic acid, the PI3K in CLL and idelalisib as well as the PDGFR mutation in hyper eosinophilia syndromes.

**Somatic (acquired) molecular abnormalities**

The molecular techniques provide advanced information in the pathogenesis of hematologic malignancies, facilitating diagnosis, treatment choice and response follow-up. The 2016 revision of the World Health Organization (WHO) classification of tumors of the hematopoietic and lymphoid tissues has accumulated new clinical, prognostic, morphologic, immunophenotypic, and genetic data, providing biomarkers that can significantly influence the diagnostic and prognostic criteria (47,48).

**Gene studies in lymphoproliferative disease: T and B cell rearrangements**

During the process of lymphocyte maturation, the rearrangements of the immunoglobulin and T cell receptor loci present a significant variety. Any clonal rearrangement refers to a hematological disorder despite some cases with viral infections [Epstein-Barr virus (EBV) or cytomegalovirus (CMV) infection]. This clonality of immunoglobulin heavy chain (IGH) and immunoglobulin kappa (IGK) loci for B cell rearrangements and T cell receptor gamma chain (TCRG) and T cell receptor beta chain (TCRB) for T cell can be detected in peripheral blood or in paraffin-fixed blocks via PCR technology, thus
permitting diagnosis of lymphoproliferative disorders. Additionally, they can potentially participate in the prognostic score and can be used as biomarkers of post treatment monitoring.

**Identification of cryptic translocation: prognostic significance**

The study of whole genome sequencing revealed the presence of balanced chromosomal abnormalities (BCAs) that are associated with disease onset and progression. Cryptic translocations and mutations in the core binding factor genes, runt-related transcription factor 1 (RUNX1) or core-binding factor subunit beta (CBFB), are identified in acute myeloid and lymphocytic leukemia, myelodysplastic syndrome and chronic myelomonocytic leukemia. The RUNX family of transcription factors (Runx1, Runx2, and Runx3) encodes proteins implicated in different cell lineages, including blood and blood-related cell lineages. They regulate via activation or repression normal hematopoietic development and they were proved valuable in distinguishing any genetic abnormality of this transcription factor. Acute promyelocytic leukemia (APL) is one of the most curable leukemias, despite its severe bleeding diathesis.

APL is characterized by a distinct chromosomal aberration, t(15;17)(q22;q21) translocation, and remains the only leukemia that shows great sensitivity to all-trans retinoic acid (ATRA) and arsenic trioxide (ATO), being treated with non-conventional chemotherapy. Genetic confirmation is essential for diagnosis. Rare cryptic fusions are often correlated with small genomic insertions and poor prognosis and they can be detected by RT-PCR although conventional chromosomal studies or even FISH analyses may appear normal (49,50).

**Mutations in acute myeloid and lymphoid leukemia**

The 2016 revision of the WHO classification of myeloid and lymphoid neoplasms is based, among others, on the representation of specific genetic mutations to add entities. ALL is a heterogeneous disorder requiring a risk-stratified approach for appropriate treatment.

Cytogenetic analysis must take place during diagnosis due to its prognostic value and molecular markers, or individual cell surface markers are used to evaluate chemotherapy response. In addition to the most frequent cytogenetic aberrations, BCR-ABL, TEL-AML1, MLL- AF4, and E2APBX1, two more have been added; B-ALL with intrachromosomal amplification of chromosome 21 (iAMP21) and BCR-ABL like B lymphoblastic leukemia/lymphoma. Moreover, AML is classified accordingly.

The molecular screening must be performed independently of the karyotype. Application of the NGS panels allows many different mutations to be tested and a better therapeutic strategy to be selected (51,52).

**MRD monitoring**

For hematological malignancies, successful treatment is considered upon induction of complete remission (CR) via chemotherapy.

The PCR techniques (RT-PCR and allele specific oligonucleotide PCR) can detect at a level $10^{-4}$ to $10^{-5}$ residual cell disease (1 leukemic cell in 10,000–100,000 healthy cells). MRD can be detected and quantified in peripheral blood or in bone marrow. Following MRD, it is helpful to organize risk groups and to monitor disease kinetics in refractory or progression cases.

MRD remains a major criterion for the decision of autologous or allogenic stem cell transplantation (SCT) (53).

**Applications to SCT**

Allogeneic hematopoietic SCT (HSCT) represents a potential curative treatment for patients with hematological malignant or non-malignant disorders. The selected donor must ideally match for human leukocyte antigens HLA-A, -B, -C, -DRB1, and -DQB1 loci.

All evolution in conditioning regimens and in the delimitation of killer immunoglobulin-like receptors (KIR) genes increases the availability of HLA-haploidentical and unrelated donors with mismatch (54,55). The donor search includes serological tests for HLA antigens and genotypic HLA typing. NGS-based methods have been refined and guarantee high resolution results. The deepest method is the allele level typing which corresponds to a unique nucleotide sequence for an HLA gene, as defined by using all digits in 1st, 2nd, 3rd and 4th fields (56,57). All those, help to decrease acute graft-versus-host disease (GVHD) and graft failure.

**Post allogeneic SCT chimerism control**

Analysis of chimerism after allogeneic SCT is necessary for graft failure detection and monitoring MRD and relapse.
The methods used are FISH (XY-FISH) analysis of sex chromosomes after transplantation from a sex-mismatched donor or short tandem repeats PCR. In order to increase sensitivity RT-PCR and highly sensitive droplet digital PCR (ddPCR) assay can be applied (58-60).

**CAR-T cell era**

Cancer genomes are highly unstable leading to diverse chromosomal abnormalities such as large insertions and deletions of chromosomal material. Karyotyping, which was for a long time the gold-standard to identify the presence of chromosomal abnormalities, suffers from the inability to identify structural abnormalities smaller than ~5 megabases. SNP and oligonucleotide microarrays have radically changed the field of cytogenetics by supplying a high resolution (a few kb) capability to identify large and small copy number variants as well as areas of copy-neutral LOH.

NGS technologies can also distinguish genomic structural variations, although routine alignment tools are not so well-equipped to perform such analyses and can only identify a few nucleotide mismatches (61-64).

**Conclusions**

Genes and proteins are the fundamental building blocks of life and have been at the heart of developments in molecular biology. We have briefly described selected developments in microfluidic technologies for molecular biology and biochemistry, with a particular emphasis on genomic and proteomic research. New molecular techniques have been applied both in diagnosis and treatment. Regarding diagnosis, the sensibility of those methods increases the precision and the speed of results while screening can be more easily performed. In respect to therapies, targeted and immunomodulatory therapies guarantee better results with less hematological toxicities. Future studies should point towards gene therapy and transplantation approaches with recombinant stem cell genes.

**Acknowledgements**

None.

**Footnote**

*Conflicts of Interest:* The authors have no conflicts of interest to declare.


46. Pui CH, Crist WM, Look AT. Biology and clinical

Cite this article as: Koutsi A, Vervesou EC. Diagnostic molecular techniques in haematology: recent advances. Ann Transl Med 2018;6(12):242. doi: 10.21037/atm.2018.05.30