Application of Molecular Karyotyping in Acute Myeloid Leukemia: A Review

Angeli Ambayya¹, Andrew Octavian Sasmita¹, Jameela Sathar¹, Kian Meng Chang¹, Zainina Seman², Sabariah Md Noor², Subramanian Yegappan¹*¹

¹Clinical Hematology Referral Laboratory, Hematology Department, Hospital Ampang, Ampang, Selangor, Malaysia
²Department of Pathology, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

Email address:
angeli_100182@yahoo.com (A. Ambayya), andrewoctavians@gmail.com (A. O. Sasmita), jsathar@hotmail.com (J. Sathar),
dechangkm@gmail.com (K. M. Chang), zainina@upm.edu.my (Z. Seman), md_subariah@upm.edu.my (S. M. Noor),
mul14141@gmail.com (S. Yegappan)
*Corresponding author

To cite this article:

Received: May 26, 2018; Accepted: May 19, 2018; Published: July 12, 2018

Abstract: Acute myeloid leukemia (AML) is an aggressive disease characterized by the overproduction of immature myeloid cells that accumulate in blood and bone marrow. Integration of genetic findings and clinicopathological information is crucial in establishing the diagnosis, prognosis and determining the therapeutic approach in the management of AML patients. In recent years, the AML classification has evolved from morphology to cytogenetics/molecular genetics-based findings, which is essential in the detection of chromosomal abnormalities and has provided the framework for the diagnosis and risk-stratification in AML. Moreover, with advances in molecular karyotyping such as comparative genomic hybridization (CGH) and single nucleotide polymorphism (SNP) arrays, various limitations of conventional diagnostic approaches have been overcome. Hence, this review focuses on the insights into molecular karyotyping using CGH and SNP arrays which enable the identification of copy number variations (CNVs) at a higher resolution and facilitate the detection of copy neutral loss of heterozygosity (CN-LOH) otherwise undetectable by conventional cytogenetics. Technical hindrances of these methods (e.g. regions of losses, gains, or “undulating waves”) are also discussed in the context of AML.

Keywords: Acute Myeloid Leukemia, Comparative Genomic Hybridization, Copy Number Variants, Genomic Analysis, Microarray

1. Introduction

Leukemia is a group of hematological malignancies characterized by abnormal proliferation of hematopoietic cells which result in impaired maturation and excessive accumulation of immature cells in bone marrow and/or blood [1]. This leads to bone marrow failure with anemia, neutropenia, thrombocytopenia or pancytopenia and increased risk of organ infiltration by the malignant cells. Leukemia itself may be broadly classified to acute and chronic leukemia. Out of 4573 cases of leukemia reported from 2007-2011 in the Malaysian population, 54.2% of the cases were AML [2], making up the majority of leukemia cases within the region. The statistics of AML within Malaysia may be compared to AML cases in the United States (US), whereby in 2015 alone, 20830 new cases were diagnosed with a recorded mortality rate of over 10000 [3]. Although treatments such as allogeneic hematopoietic stem cell transplantation (HSCT) [4] and induction therapy with cytarabine [5] do exist to improve overall survival (OS) of patients, untreated patients over the age of 60 are reported to only have an average survival between 5 to 10 months [6], further emphasizing the importance of a robust diagnostic approach to better stratify the disease. Thus, this review focuses on the deoxyribonucleic acid (DNA) microarray-based molecular cytogenetic approaches, advances, and
limitations in the context of AML with the addition of how these may be superior to conventional methods to diagnose this disease.

2. Overview on Current AML Diagnostic Approaches

Over the years, the primary diagnosis of AML lies on the morphologic assessment of leukemic myeloblast in the peripheral blood or bone marrow. Diagnosis of AML is established in the presence of blast exceeding 20% in the blood or marrow with the exception for cases with t (15; 17), t (8; 21), t (16; 16) or inv (16) and erythroleukemia. In some part of the world, lineage assignments depend on cytochemistry such as myeloperoxidase (MPO) or Sudan Black B (SBB), non-specific esterase (NSE) and periodic acid Schiff (PAS) [7, 8]. Morphological assessment, however, has limitations as it requires expertise to distinguish the cells, especially in cases where ambiguities exist in differentiating leukemic cells from reactive cells, whereby other ancillary testing is required to confirm the lineage of the leukemic cells.

Aside from morphologic assessment, flow cytometry immunophenotyping (IP) is an imperative tool in the diagnosis of AML as it enables the characterization of leukemic cells through assessment of cell surface, cytoplasmic and nuclear antigen expression patterns [9–11]. Characteristic expressions of myeloid lineage markers such as CD13, CD33 and CD117 enable the distinction of AML from other types of leukemia. Asynchronous antigen expression, also known as aberrant antigen expressed in leukemic cells serves as an auxiliary tool in the diagnosis and prognosis of AML. Aberrant expressions of CD56 in AML and CD2 in APL have been implicated as poor prognosis. In particular, CD56 expression in AML with t (8;21) corresponds to a shortened DFS and OS, including in patients who have undergone transplantation [12]. In addition, flow cytometry immunophenotyping (IP) also aids in the monitoring of minimal residual disease (MRD) with greater sensitivity as compared to morphological examination of bone marrow aspirates, though at a 1-log less sensitive as compared to polymerase chain reaction (PCR) based techniques [13]. Moreover, MRD by flow cytometry IP is possible for cases where molecular techniques cannot be utilized. However, there are drawbacks in this technique especially phenotypic shifts during and/or after therapy [14]. Secondly, there is no clear consensus on the list of markers for AML studies and so the selection of markers is usually based on the individual laboratory’s discretion and experience. However, with larger adoption of multicolour flow cytometry, this drawback will likely be mitigated in the future.

Cytogenetics remains an indispensable tool in prognostication in AML and plays a key role in the clinical management of AML patients. Specific cytogenetic aberrations are associated with specific abnormalities that are subsequently correlated with response to therapy and patient’s survival. However, efforts to elucidate sub-microscopic genetic alterations are thwarted by limited resolution, poor chromosome morphology aside from other sampling and technical issues [15]. Techniques such as FISH are useful in identification of structural and numerical aberrations as it targets non-dividing cells (interphase FISH). This technique, however, is more of a targeted approach and requires prior knowledge of anomalies of interest; thus, it is unsuitable as a screening tool in the diagnosis of leukemia.

Based on the conventional cytogenetics method, the World Health Organization (WHO) reported a classification of AML in 2016, whereby the disease was subdivided into 25 entities with additional subclasses: provisional entities where mutations occurring in AML with normal karyotype namely NPM1 and CEBPA; AML associated with Down syndrome, acute panmyelosis with myelofibrosis, granulocytic sarcoma, and blastic plasmacytoid dendritic cell neoplasm as shown in Table 1 [16]. Having been shown to possess different treatment responses, this classification helps to improve prognosis of AML through molecular testing and further stratify predicted patient outcomes. Dohner et al. for instance, further associated molecular signatures that some AML sub-classes have into favorable, intermediate I, intermediate II, and adverse genetic groups [7].

<table>
<thead>
<tr>
<th>Table 1. WHO classification of AML and related neoplasms (2016) [16].</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AML with recurrent genetic abnormalities</strong></td>
</tr>
<tr>
<td>AML with t(8;21)(q22;q22.1);RUNX1-RUNX1T1</td>
</tr>
<tr>
<td>AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22);CBFB-MYH11</td>
</tr>
<tr>
<td>AML with PML-RARA</td>
</tr>
<tr>
<td>AML with t(9;11)(p21.3;q23.3);MLT3-KMT2A</td>
</tr>
<tr>
<td>AML with t(6;9)(p23;q34.1);DEK-NUP214</td>
</tr>
<tr>
<td>AML with inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2);GATA2,COM</td>
</tr>
<tr>
<td>AML (megakaryoblastic) with t(1;22)(p13.3;q13.3);RBM15-MKL1</td>
</tr>
<tr>
<td>Provisional entity: AML with BCR-ABL1</td>
</tr>
<tr>
<td>AML with mutated NPM1</td>
</tr>
<tr>
<td>AML with biallelic mutations of CEBPA</td>
</tr>
<tr>
<td>Provisional entity: AML with mutated RUNX1</td>
</tr>
<tr>
<td>AML with myelodysplasia-related changes</td>
</tr>
<tr>
<td>Therapy-related myeloid neoplasms</td>
</tr>
<tr>
<td>AML, NOS</td>
</tr>
<tr>
<td>AML with minimal differentiation</td>
</tr>
<tr>
<td>AML without maturation</td>
</tr>
<tr>
<td>AML with maturation</td>
</tr>
<tr>
<td>Acute myelomonocytic leukemia</td>
</tr>
<tr>
<td>Acute monoblastic/myelocytic leukemia</td>
</tr>
<tr>
<td>Pure erythroid leukemia</td>
</tr>
<tr>
<td>Acute megakaryoblastic leukemia</td>
</tr>
<tr>
<td>Acute basophilic leukemia</td>
</tr>
<tr>
<td>Acute panmyelosis with myelofibrosis</td>
</tr>
<tr>
<td>Myeloid sarcoma</td>
</tr>
<tr>
<td>Myeloid proliferations related to Down syndrome</td>
</tr>
<tr>
<td>Transient abnormal myeloplasia (TAM)</td>
</tr>
<tr>
<td>Myeloid leukemia associated with Down syndrome</td>
</tr>
</tbody>
</table>

The success of conventional cytogenetics analysis is largely dependent on number of mitotic cells but in some cases, obtaining enough number of mitotic cells is difficult. However, these limitations have been largely overcome by the development of molecular karyotyping which are not
affected by cell cycles and viability of leukemic cells as the starting material for the experiment is DNA, not cell cultures. In addition, molecular karyotyping tools such as array CGH (aCGH) and SNP arrays (SNP-A) enabled the detection of previously unrecognized chromosomal lesions. In contrast to aCGH which only detects the copy number changes at a high resolution, SNP-A allows both; detection of copy number changes and SNP-based genotype at sub-megabase resolution unmasking small areas of copy number loss of heterozygosity (CN-LOH), also referred to as uniparental disomy (UPD).

3. Genome Wide Analysis Using Array-Based Techniques

Genome wide analysis for chromosomal rearrangements and imbalances using array based techniques holds great appeal, as it offers interrogation of the AML genome with unprecedented resolution and supplants some of the shortcomings of conventional cytogenetics. Numerous array-based technologies have been developed and optimized for a wide range of applications in molecular genetics. In the beginning, array based technologies were first used for gene expression profiling and then followed by CNV studies (array CGH) and SNP genotyping. Currently these technologies have also been implemented in DNA methylation, alternative splicing, micro RNA (miRNA) and protein-DNA interactions (array based Chromatin ImmunoPrecipitation; ChIP) [17].

3.1. Comparative Genomic Hybridization (CGH)

Array CGH makes detection of CNVs on a genome wide scale with higher resolution possible. Array CGH is suitable for the analysis of cells, tissues, and even formalin fixed paraffin embedded (FFPE) samples without the need of culturing. Earlier aCGH platforms were based on large insert clones such as bacterial artificial chromosomes (BAC), yeast artificial chromosome (YAC) or P1-derived artificial chromosomes (PAC) clones [18]. However, producing BAC clones for array experiments is expensive and time consuming and the limits of BAC aCGH have been reached. On the other hand, oligo aCGH is flexible in probe design, has greater coverage and much higher resolution. Oligonucleotide typically contains single stranded 25-85 mer on the array [19]. The technique of labeling and hybridization of oligo aCGH differs according to the platform.

Array CGH aids in CNVs detection by evaluating signal intensities of two differentially labeled DNA (test and reference DNA) which are co-hybridized to a sequence of probes usually representing the entire genomic regions immobilized on an array slide. Generally, isolated genomic DNA (blood or other tissues) are differentially labeled by random priming where the experimental DNA will be labeled with Cyanine 5-dUTP (Cy5) dye and the reference DNA of a known genotype will be labeled with Cyanine 3-dUTP (Cy3) dye [20] (Figure 1) After hybridization, the amount of fluorescence is determined by comparing the log ratio of experimental DNA versus reference DNA. A negative loss ratio indicates CN losses and a positive log ratio corresponds to a CN gain.

3.2. Single Nucleotide Polymorphism (SNP) Array

In contrast to array CGH which only permits the identification of CNVs, SNP array facilitates simultaneous evaluation of DNA copy number alterations and genotyping information for detection of CN-LOH [21]. Prior to interrogation of genomic loci, DNA is subjected to restriction digestion and the fragmented DNA is hybridized on a slide which contains oligonucleotide probes. Measurement of the copy number of each allele for a given SNP is compared to a known reference. The signal intensity is converted into two types of output: genotype for the individual SNP (A, B or AB as the canonical genotype) and copy number values. Usually, heterozygosity is identified as “AB” and homozygosity as either “A” or “B” at each SNP locus. Detection of CN-LOH can be established by identifying regions of homozygous constellation with diploid copy number (Figure 2).

Loss of heterozygosity can be defined as sporadic loss of whole or partial segments of one of the two parental chromosome homologs. Hemizygous loss is segmental loss of one homolog while the other copy remains intact. A change from a heterozygous state to homozygous state can thus be defined as LOH [22].

On the other hand, CN-LOH happens when a genomic region originated from one parent and the loss of the other
concurrent allele [21][23]. Duplication of maternal (unimaternai) and loss of paternal and vice versa, results in CN-LOH without affecting the diploid copy number. CN-LOH also often termed as UPD, can be either constitutional or acquired affecting either whole or a segment of a chromosome.

4. CGH + SNP in a Single Array

Recently, SNPs have been added to array CGH which enable elucidation of CNVs and CN-LOH. This provides insights into regions of acquired homozygosity that may harbor tumorigenic changes that possibly predispose to AML. Combined arrays have shortened sample processing time as the protocol of CGH and SNP are combined into a single experiment which leads to a less laborious workflow. Data analysis becomes feasible as dual data of CGH and SNP is presented concurrently. Although SNP probe coverage on the combined platform tends to be lower than the traditional platform, other studies have highlighted that the performance of this combined array is exceptional in aiding investigation of CN-LOH, mosaic aneuploidy and CNVs [24].

5. Mosaicism in AML

In mosaic samples, there will be proportion of normal cells and tumor cells resulting in a fractional of total CN in non-diploid regions. Previous studies have theorized the presence of mosaicism by assessing log2 ratios or allele specific copy number and predicate the differences as potential indicators of clonal heterogeneity [25-27]. In a study on clonal evolution in hematologic malignancies by combining comparative genomic hybridization and single nucleotide polymorphism array on 16 patients, manual peak reassignment analysis enabled the detection of heterogeneity in 10 patients [27]. In this study, the researchers identified the clonal heterogeneity by referring to clonal fraction, discrepancy between CGH and SNP data, distribution of log2 ratios and peaks in the distribution plot where a CN was not assigned as shown in Figure 3.

Figure 2. SNP array workflow and interpretation. Probes in a SNP array experiment are designed to include restriction digestion sites of AluI and RsaI. Both cut alleles in the target DNA (AA, uncut allele) will result in low signal and for both uncut alleles (BB, 2 uncut alleles) will result in high signal. For one uncut allele (AB, 1 uncut allele) the signal will be intermediate. Copy number of one allele at each SNP is compared against a known reference. Genotype information for normal (diploid), loss, amplification and CN-LOH is shown.

Figure 3. Clonal heterogeneity indicators. A) View of chromosomes generated by Cytogenomics software version 2.9.2.4 which shows a deletion in the p arm of chromosome 12. Picture B) and C) show the CGH and SNP data of the chromosome 12. Different log2 ratios indicate the presence of sub-clones in the sample. Clonal fraction based on log2 ratios were estimated to be 79.3% (A) and 47.8% (B) respectively. CGH data shows a deletion of 12p but SNP data was normal for chromosome 12. This indicates a presence of sub-clone in the sample. Manual peak reassignment was performed and the SNP data showed similar findings with CGH data post reassignment (C).
In our study on AML genome, clonal heterogeneity was by referring to clonal fraction discrepancy in our customized combined array CGH+SNP data by manual peak reassignment analysis. Low amplitude of array CGH was observed without confirmation of SNP probes which indicated presence of minor clone. Figure 3 displays data before and after manual peak reassignment in an AML patient.

6. Limitations of SNP Array and CGH Array

The American College of Medical Genetics and Genomics (ACMG) has outlined several limitations of DNA microarray analysis that include: inability to detect balanced chromosomal translocation and inversions [28]. However, this has been overcome by the introduction of translocation CGH (tCGH) which not only spots the breakpoints but, more importantly singles out the translocation partner. Secondly, the presence of low tumor cells admixed with normal cells reduces the ability to detect tumor-related changes. However, Valli et al. reported as low as 8% of abnormal cells in mosaic samples [26]. Array CGH is unable to identify tetraploidy and other ploidy levels [29] [30] but a combined array CGH+SNP platform is capable in revealing polyploidy. Other than that, ACMG also highlighted the inability of these platforms to discriminate clonal and sub-clonal populations [28]. Nevertheless, log2 ratio may provide a hint of clonal heterogeneity. Software which are equipped with tools to intervene and reassign the copy numbers also exist for cases that have major and minor clones.

7. Evaluation of Tumor DNA in Microarray Analysis

The evaluation of tumor sample along with matched normal genomic DNA is recommended by the American College of Medical Genetics (ACMG) Standards and Guidelines and other researchers [24] [31] [32] [33]. Germline DNA sources vary in different studies: buccal swabs, skin biopsies, bone marrow or peripheral blood samples collected when patients are in remission after induction, saliva and urine.

In a study by Rasi et al., saliva DNA was evaluated as matched germline in chronic lymphocytic leukemia on genome wide SNP arrays. This study also reported that urine provided similar quality of germline DNA but in a lower amount. Rasi et al. also suggested that saliva represents a practical source of germline DNA for genomic hematological malignancies studies [34]. Walter et al (2009) studied 86 de novo adult AML using 1.85 million feature SNP arrays with paired normal derived from normal tissue (skin biopsy) and tumor DNA analysis and delineated 201 somatic CNVs [35]. Similarly, Tiu et al (2011) studied myelodysplastic syndromes and related myeloid malignancies by using paired bone marrow and germline (CD3+ cells) to distinguish germline lesions [36]. Although can be obtained easily, usage of buccal swab at diagnosis in hematologic malignancies may be contaminated with tumor cells due to the presence of high white cell counts and circulating blast whereas skin biopsy is not routinely done in the study sites [37] [38]. On the other hand, using remission after induction samples is feasible as the morphologic features and presence of tumor cells can be assessed to consider the suitability of the selected sample to be used as the source of germline DNA prior to DNA extraction. However, obtaining a germline sample from patient with refractory or relapsed disease will be a significant limitation when remission after induction sample is used as the method of choice to obtain germline DNA.

Conversely in other studies, leukemia related CNVs were reported by referring to CNV databases [39-41]. Although CNV databases such as The Database of Genomic Variants (DGV) are useful in delineating germline variants, there is probability of excluding somatic alterations related to leukemia leading to underestimation of the actual percentage of genomic aberrations [42]. Therefore, matched normal genomic DNA is the best reference for leukemia studies [24, 31, 33].

In our study, paired tumor and germline (remission after induction sample obtained from the same patient) DNA were used to delineate germline variants. Remission sample were evaluated by morphologic assessment of peripheral blood smears where the blast count was below 5%, followed by assessment by flow cytometry immunophenotyping for Minimal Residual Disease (MRD) detection where the leukemic phenotypic profiles seen at diagnosis were not detected. A stringent algorithm was applied in identifying inherited sequence variants unrelated to the tumor. Tumor and matched germline DNA aberrations and LOH intervals reports generated by Cytogenomics software version 2.9.2.4 were compared for all cases. Matching aberrations present in tumor and matched germline DNA were considered germline and changes present only in tumor were classified as tumor related changes. Overestimation of genomic aberrations was avoided by germline DNA comparison. In our study, only 22% of the tumor related aberration where CN-LOHs but in germline DNA, about 87% of genomic aberrations were CN-LOH. This proves that it is critical to use the corresponding germline DNA in each of the AML cases as large number of CNVs and CN-LOHs were present as inherited sequence variants and appear to be unique to each individual patient.

Apart from using the germline DNA to delineate inherited sequence variants, a database of CNVs derived from “normal” multi-ethnic population of Malaysia was developed. From this, 353 novel and recurrent CNVs were made available as customized track in the array CGH+SNP analysis software as shown in Figure 4. This provides invaluable information in the analysis of Malaysian AML patients apart from utilizing other databases such as DGV.
With regards to analysis, studies need to determine the CNVs interval sizes, pathogenic aberrations as well as the diagnostic, prognostic and therapeutic significance of the abnormalities detected. Further studies on clinically uncertain, acquired and likely pathogenic somatic aberrations should be carried out prior to reporting of these abnormalities to the clinicians without speculating the pathogenicity of the CNVs.

8. Studies on AML Using Array CGH and SNP Array

During the last decade, tremendous advances have been witnessed in the field of molecular karyotyping, particularly the application of array CGH and SNP array. Incorporation of clinicopathologic and molecular genetics discovery not only aids in the diagnosis but also in predicting AML outcome.

Notable studies on normal karyotype AML (AML-NK) have been successful in unmasking cryptic genomic aberrations previously undetected by conventional cytogenetics. A study by Chin et al. on cytogenetic profile of de novo AML patients in Malaysia disclosed that about 69.6% of 480 cases showed normal karyotype by conventional cytogenetic analysis [43]. Studies from other countries reported lower frequency of AML-NK of 40-50% [39, 44, 45]. The high percentages of AML-NK may be due to resolution limitations of conventional cytogenetics and other confounding sample and technical related issues. With the adoption of array CGH and SNP array, these limitations have been overcome in some of the studies of not only AML-NK, but various subtypes of AML [46] [47].

A study by Akagi et al (2009) on 30 AML-NK and 8 MDS cases using GeneChip Human mapping 250K array Nspl microarray (SNP-chip, Affymetrix, Santa Clara, CA, USA) showed that 51% of the cases had a normal karyotype by SNP array analysis [39]. On the other hand, 49% of the samples displayed a minimum of one or more genomic aberrations including gain, loss and CN-LOH. Another study using Genome-Wide Human SNP Array 6.0 (Affymetrix, Santa Clara, CA) was carried out on AML-NK with a larger number of subjects (n=133) by Jun et al (2011) which revealed 67.7% had normal karyotype by SNP array analysis. In contrast 32.3% of the patients had at least one abnormal SNP lesion that was undetected by the conventional cytogenetics method [48]. A total of 113 aberrations were detected, which included 23 gains, 55 losses and 35 CN-LOH.

Copy number variations using array CGH were also reported in diverse types of AML cases. Rucker et al. studied...
60 AML with complex karyotypes using a custom designed microarray consisting 2,799 BAC aCGH and PAC aCGH clones with parallel analysis of gene expression in selected cases [49]. This study reported that the frequency of genomic losses was higher than gains. The regions with highest frequencies of losses were 5q (77%), 17p (55%) and 7q (45%) and the most frequent gains were noted in 11q (40%) and 8q (38%). Candidate genes which may be involved in the pathogenesis of AML were identified by combined analysis of array CGH and gene expression profiling.

Recently Zhang et al. focused their study on AML-M5 using array CGH consisting about 720,000-mer oligonucleotide covering coding and non-coding regions [40]. This study elucidated new AML-M5 associated CNVs, including gains of 3q26.2-qter and 13q31.3 as well as losses of 2q24.2, 8p12 and 14q32. Clinically significant recurrent deletion of 14q32 (FOXXN3 gene) were singled out from this study. However, FOXXN3 gene deletion is not specific for AML-M5 as it also reported in other myeloid leukemia. Zhang et al. also reported amplification of MLL and deletion of CDKN2A which were seen concurrently with potential link to AML-M5.

In another study by Mehrotra et al., 48 AML uniformly treated AML patients were tested with a custom-designed 4X44K, 60-mer oligonucleotide genomic array with gene-centric whole genome coverage [50]. Mehrotra’s research group detected a total of 170 CNVs and reported that frequencies of sub-microscopic aberrations were higher compared to whole chromosomal abnormalities. Array CGH also enabled the discovery of previously underappreciated cryptic aberrations as small as 5 kb, although the functional significance yet to be known. Additionally, array CGH enabled the delineation of chromosomal breakpoints in del (5q) which ranged from position 5q11.2 to 5q34. Interestingly, this study associated the chromosomal aberrations with patient’s OS and CR status. Recurrent loss of a 3.2 Mb in region of 17p11.2-p13 were consistent with a shorter OS with a median of 34 weeks in patients. On the other hand, the loss of 155 kb in the region of 5q33.3 was associated with CR.

In our study, paired tumor and germline DNA (remission sample obtained from the same patient after induction) analysis was performed using a customized Agilent CGH+SNP 180K DNA microarray with additional custom probes for 49 genes at exon resolution on 41 AML patients [51]. A total of 1191 aberrations in the germline and tumor DNA ranging from 0.000224 Mb to 138.20 Mb were detected. In depth study revealed that of the 1191 aberrations; 449 were somatic, 386 lesions were germline (present in tumor and germline DNA) and 356 were inherited sequence variants seen only in the germline DNA. Somatic lesions were grouped based on sizes (n=1 Mb=213, n 1-5 Mb=105, n 5-10 Mb=20, n 10-20Mb=7, n >20 Mb=19). Gains were most frequent in chromosome 19 (19/250 cases) whereas losses were most frequent in X chromosome (15/99 cases). CN-LOHs were frequently seen in chromosome 4 (9/15 cases). Thirty-one recurrent aberrations were identified in this study. Most recurrent aberrations were gains in chromosome 19q13.2 (5 cases), chromosome 10q24.32 and 1q43 (4 cases respectively). Somatic aberrations were observed in three of the customized exon resolution genes: ATM (gains in 3 cases, losses in 2 cases), EZH2 (gain and loss in one case respectively), and MLL (gains in 2 cases). Interestingly, a higher level of aberrations was observed in the telomeric regions of tumor DNA (31.0%) compared to germline DNA (12.7%). Most of these aberrations were gains; tumor (88.5%) and germline (91.8%).

In another study, 53 genomes of AML patients were studied using the same design of customized array CGH+SNP DNA microarray [52]. Patients with complete remission (CR) following induction therapy were compared with those who expired either before or just after induction chemotherapy. A total of 449 somatic aberrations in patients in CR group (ranging from 0.23 kb to 138.20 Mb) were identified and 530 aberrations (range 1.58 kb to 147.67 Mb) were seen in the expired patient group. Statistical analysis revealed 4 significant novel regions of losses seen exclusively in the expired patients that could be related to their poor OS (P < 0.01 by Fisher’s exact test): 1q11.22, 8p23.1, 22q11.21 and 8p11.22. On the other hand, recurrent gains seen in 19q13.2, 10q24.32 and 1q43 showed significant association with patients in CR suggesting a predictor of better OS in AML patients (P < 0.01 by Fisher’s exact test). Within the analysis, 29 losses and 58 gains were identified which involved 28 genes that have customized probes for every exon. Mutually exclusive exonic CNVs in the genes were discovered between the two groups (7 genes in the CR group and 11 genes in the expired group). Recurrent CNVs observed on the exons of TP53 gene in the expired patients and TET2 in patients in CR.

Summarized in Table 2 are comparisons between various studies considering the AML genome using different types of array CGH and SNP platforms. Our analysis also revealed high burden of CNVs in our cases, which contradict previous reports [35] [53]. However, these findings significantly support reports by Mehrotra et al. and Jun et al. whereby the studies observed more than two-fold of CNVs in their samples [48] [50]. Studies by Mehrotra et al. and Jun et al., however, lacked systematic use of germline DNA and thus, there could be overestimation of somatic CNVs. Systematic use of matched tumor and germline samples enabled unequivocal delineation of somatic genetic alteration including CNVs and CN-LOHs in our study. Therefore, germline CNVs and CN-LOHs were excluded from our analysis leading to lower numbers of genomic aberrations in our study.

Based on these findings, undeniably array CGH and SNP array have unprecedented advantages in characterizing chromosomal aberrations on a genome wide scale. These genome wide scanning tools facilitate the discovery of a plethora of genomic alterations in AML which may lead to improved diagnosis and prognostic stratification of patients and hold promises for innovation of targeted therapy in the future.
Table 2. Comparison of findings from various studies involving AML on genomic microarray platforms.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>38</td>
<td>86</td>
<td>133</td>
<td>48</td>
<td>41</td>
</tr>
<tr>
<td>Array type</td>
<td>250K SNP</td>
<td>SNP 6.0</td>
<td>SNP 6.0</td>
<td>Custom 44K oligonucleotide genomic array</td>
<td>Custom CGH+SNP 180K</td>
</tr>
<tr>
<td>Cases</td>
<td>30 AML-NK, 8 MDS</td>
<td>60 de novo AML</td>
<td>133 AML-NK</td>
<td>AML</td>
<td>41 AML</td>
</tr>
<tr>
<td>Use of Germline</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CNVs</td>
<td>19 cases had one or more CNVs</td>
<td>201 CNVs</td>
<td>Total 113 SNP lesions</td>
<td>170 total chromosomal aberrations (including microdeletion)</td>
<td>1191 aberrations (449 somatic; 386 lesions in both tumor and germline DNA; 356 germline only)</td>
</tr>
<tr>
<td>Gains</td>
<td>Deletion and duplication in 9 patients</td>
<td>201 CNVs</td>
<td>23</td>
<td>71</td>
<td>19</td>
</tr>
<tr>
<td>Loss</td>
<td>-</td>
<td>-</td>
<td>55</td>
<td>99</td>
<td>15</td>
</tr>
<tr>
<td>CN-LOH</td>
<td>12 patients</td>
<td>8 regions of UPD in 7 of 86 samples.</td>
<td>35</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>Normal karyotype by CGH and/or SNP</td>
<td>19 patients</td>
<td>43</td>
<td>90</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>One or more genomic aberrations by CGH and/or SNP</td>
<td>18 patients</td>
<td>38</td>
<td>43</td>
<td>27</td>
<td>19</td>
</tr>
</tbody>
</table>

9. Technical Considerations in Array CGH and SNP Findings

During the last decade, tremendous advances have been witnessed in the field of molecular karyotyping, particularly the application of array CGH and SNP array. Incorporation of clinicopathologic and molecular genetics discovery not only aids in the diagnosis but also in predicting AML outcome.

Notable studies on normal karyotype AML (AML-NK) have been successful in unmasking cryptic genomic aberrations previously undetected by conventional cytogenetics. A study by Chin et al. on cytogenetic profile of de novo AML patients in Malaysia disclosed that about 69.6% of 480 cases showed normal karyotype by conventional cytogenetic analysis [43]. Studies from other countries reported lower frequency of AML-NK of 40-50% [39, 44, 45]. The high percentages of AML-NK may be due to resolution limitations of conventional cytogenetics and other confounding sample and technical related issues. With the adoption of array CGH and SNP array, these limitations have been overcome in some of the studies of not only AML-NK, but various subtypes of AML [46] [47].

9.1. Gains in the Telomeric Regions of Chromosome 14q32.33

There were ambiguous gains in chromosome 14q32.33 in 11 cases extending to the telomeric regions as shown in Figure 4. A review of these cases showed that there were insufficient amounts of CGH and SNP probes in this region to support the findings. A review on the reference DNA (Figure 5A) showed that the reference DNA (male) has a deletion. Signal intensity plots revealed that the gain in the sample S29 was most probably due to the deletion in the reference DNA where the values of the Cy3 (reference DNA) are below 1000. Typically, signal intensity ratio of a 2n copy number of reference DNA to a sample DNA should be 1000 after normalization. An arbitrary value of 500 indicates single copy deletion and a value of below 500 may indicate two copy deletions. These cases were excluded from subsequent analyses.

9.2. Gains in the Telomeric Regions of Chromosome 14q32.33

Similarly, equivocal genomic centromeric aberrations were seen in a total of 17 cases (3 loss, 14 gain) in chromosome 9p13.1-p11.2; ranging from 4.53 – 8.75 Mb. These findings were excluded from our analysis as the coordinates of the aberrations were similar in 12 cases which indicated that these may not reflect true aberration and sizes but the terminal of 9p centromeric region. Moreover, there are no SNP probes in this region to reconfirm the findings as shown in Figure 5B. should be 1000 after normalization. An arbitrary value of 500 indicates single copy deletion and a value of below 500 may indicate two copy deletions. These cases were excluded from subsequent analyses.

9.3. Region of Undulating “Waves” in Chromosome 16q21

Reproducible regions of equal copy numbers were also identified with a small increase in the average CGH log ratio (AvgCGHLR) varying from 0.151-0.168 in eight cases (7 tumor samples, 1 germline) in chromosome 16q21. These cases are shown in Figure 5C.
10. Conclusion

AML is a heterogeneous hematological disorder which progresses rapidly and requires immediate treatment. It requires incorporation of various clinical and laboratory findings in a timely manner to treat the patients effectively. Conventional cytogenetics has been extremely valuable in prognostication of AML over the last three decades. However, there are limitations in this method where it is laborious in nature, requires technical expertise and thwarted by inability to uncover many potential sub-microscopic genomic alterations due to resolution limitations.

With the use of combined array CGH+SNP platform, a comprehensive approach for elucidation of clinically relevant CNVs and CN-LOH were possible in a single assay. Delineation of tumor related genomic aberrations from inherited genomic variants was possible with the use of matched germline DNA for each of the case and minimized false discoveries in our case series. Chromosomal
abnormalities with breakpoints coordinates could be more accurately compared to conventional cytogenetics. Array CGH+SNP had several distinct advantages in terms of resolution and detection of cryptic CNVs aside from regions of CN-LOH which were undetectable by conventional cytogenetics. Besides that, in cases with no analyzable metaphase using conventional cytogenetics, array CGH+SNP were useful in providing chromosomal aberrations as no cell culturing is required as the starting material was DNA. These findings can then be confirmed with other methods such as FISH, quantitative PCR (qPCR), multiplex ligation-dependent probe amplification (MLPA), or a different microarray platform.

In terms of analysis, the laboratory needs to determine the CNVs interval sizes, pathogenic aberrations as well as the diagnostic, prognostic and therapeutic significance of the abnormalities detected. Further studies on clinically uncertain, acquired and likely pathogenic somatic aberrations should be carried out prior to reporting of these abnormalities to the clinicians without speculating the pathogenicity of the CNVs.

Although there are rapid advancements in the other molecular techniques, array CGH and SNP holds great potential to be integrated in clinical management of AML patients in terms providing diagnosis, prognostication and stratification of patients according to their genomic abnormalities. Ultimately this will pave the way to development of enhanced treatment modalities as this technique provides faster detection of chromosomal aberrations with greater accuracy in mapping the breakpoints in somatic aberrations detected in AML.

Acknowledgements

Firstly, we would like to thank God Almighty for His grace. We extend our gratitude to the Director General, Ministry of Health, Malaysia for permission to publish and the Ministry of Health, Malaysia for funding our preceding research which supplements this review. Lastly, we would also like to acknowledge the following: Phan Chin Lee, Carey Wee, Daniele Belluoccio and the staffs of Clinical Hematology Referral Laboratory for their help and support.

References


29 Angeli Ambaya et al.: Application of Molecular Karyotyping in Acute Myeloid Leukemia: A Review


