

A Pilot Study of Next-Generation Sequencing on Cell-Free DNA from Blood Plasma and Bone Marrow Fluid for Detecting Leukemic Clonal Abnormalities

Qinghua Zhou, PhD;^{1*} Taiwo Togun, PhD;² Peining Li, PhD;¹ Zongzhi Liu, PhD^{2*}

¹ Department of Genetics, Yale University School of Medicine, New Haven, CT

² SeqWise NGS Counseling Inc, Orange, CT

Next-generation sequencing on maternal plasma cell-free DNA (cfDNA) has been applied to non-invasive prenatal screening for common aneuploidy in the fetuses. It has been proposed that cfDNA could be a useful biomarker for early cancer detection, residual disease discovery, and chemotherapy monitoring. In patients with hematological malignancies, the cancerous cells undergoing apoptosis could release leukemic cfDNA into the blood plasma or bone marrow fluid, and the chromosomal profiling from those cfDNA could be used to detect clonal chromosome abnormalities. To evaluate the technical and clinical feasibility of next-generation sequencing on cfDNA for detecting leukemic clonal abnormalities, a pilot study was performed on ten residual samples to compare results from cfDNA sequencing analysis (cfDSA) with diagnostic findings from karyotyping, fluorescence in situ hybridization (FISH) and array comparative genomic hybridization (aCGH). Six of the ten samples had normal karyotypes, and consistently, both cfDSA and aCGH showed normal results. In three samples with different clonal chromosome abnormalities, aCGH and cfDSA detected comparable copy number aberrations and further defined the chromosomal abnormalities. In one case with FISH-detected deletions of 7q and 20q in 11-12% of cells, neither cfDSA nor aCGH detected any copy number aberrations. The result from this pilot study demonstrated that leukemic cfDNA in the blood plasma, and possibly bone marrow fluid, could be used to detect clonal chromosome abnormalities. However, the analytic and clinical validities need to be established using a large sample series and user-friendly designed bioinformatic tools need to be developed for robust sequencing data analysis in a clinical setting.

[*N A J Med Sci.* 2014;7(4):180-185. DOI: 10.7156/najms.2014.0704180]

Key Words: Leukemia, cell-free DNA (cfDNA), next-generation sequencing, cytogenetics, clonal abnormality

INTRODUCTION

Leukemia is a disease with clonal proliferation of cancer hematopoietic stem cells in the bone marrow. Overall, over 50% of leukemic cases contain chromosomal abnormalities.¹⁻³ Current cell-based cytogenetic analysis includes karyotyping for detecting clonal chromosome abnormalities and fluorescence *in situ* hybridization (FISH) for identifying recurrent gene rearrangements in different types of leukemia.⁴ DNA-based array comparative genomic hybridization (aCGH) has been introduced as a supplemental test to define chromosomal and submicroscopic imbalances for chronic lymphocytic leukemia (CLL), acute lymphocytic leukemia (ALL), myelodysplastic syndrome (MDS), and acute myeloid leukemia (AML).^{1,2,5}

Next-generation sequencing on maternal plasma cell-free DNA (cfDNA) has been applied to non-invasive prenatal screening for fetal aneuploidy since 2008.⁶⁻⁸ The cfDNA in the blood plasma or serum was considered as an ideal biomarker for detecting various types of cancers or monitoring cancer treatment,⁹⁻¹³ as it provides the possibility of avoiding tumor tissue biopsies and tracing the changes in cfDNA during the natural course of the disease or following cancer treatment. For different types of cancer, analyzing cfDNA could serve as a “non-invasive biopsy” for numerous diagnostic applications, such as early cancer detection,^{10,12-13} residual disease discovery,⁹ or chemotherapy monitoring.¹¹ However, cfDNA-based analysis for cancer is still in the stage of basic research and few attempts have been made to translate it into a clinical diagnosis. In leukemia patients, a portion of cfDNA in the blood plasma or bone marrow fluid is originated from the apoptosis of cancer cells. Apoptotic cells are usually phagocytosed by macrophages, which can release digested DNA into the tissue environment.⁹ It has been noted that cfDNA from leukemic cells has higher integrity than that from healthy cells. Therefore, the genomic

Received: 09/02/2014; Revised: 09/29/2014; Accepted: 10/01/2014

*Corresponding Authors: Qinghua Zhou, PhD, Department of Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520-8005. Tel: 203-785-6317; Fax: 203-785-7342.

(Email: qinghua.zhou@yale.edu)

Zongzhi Liu, PhD, SeqWise NGS Consulting Inc, 274 Hemlock Drive, Orange, CT 06477. Tel: 203-903-2599; Website: www.seqwise.com.

(Email: zongzhi.liu@seqwise.com)

information from leukemic cfDNA could reflect the clonal abnormalities of cancer cells.¹⁴ A pilot study was performed to evaluate the technical and clinical feasibility through a parallel comparison between cytogenetic and cfDNA sequencing findings. The results prompt further analytic and clinical validation and better bioinformatics tools for cfDNA sequencing analysis (cfDSA).

MATERIALS AND METHODS

Diagnostic Cytogenomic Analyses

Yale Molecular Cytogenetics and Genomics Laboratory routinely performs cytogenomic analyses for various types of leukemia. Chromosome analysis was performed following

laboratory's standardized procedures; FISH tests were conducted using probes for leukemia-specific panels of recurrent gene rearrangements (Rainbow Scientific, Inc., Windsor, CT). The FISH probes include *DLEU1* at 13q14.1, *DI3S1825* at 13q34, *DI2Z3* at 12q15, *RELN* at 7q22.1, *TES* at 7q31.2, *MYC* at 8q24.21, and *PTPRT* at 20q12. Genomic DNA was extracted from blood leukocytes or bone marrow cells using the Puregene kit (Gentra systems, Inc. Minneapolis, MN). The aCGH using 8x60K CGH array (Agilent Technologies, Inc., Santa Clara, CA) was performed following manufacturer's manual, and the copy number aberrations were described using NCBI36/hg18 assembly of Human Genome Browser (<http://genome.ucsc.edu/>).

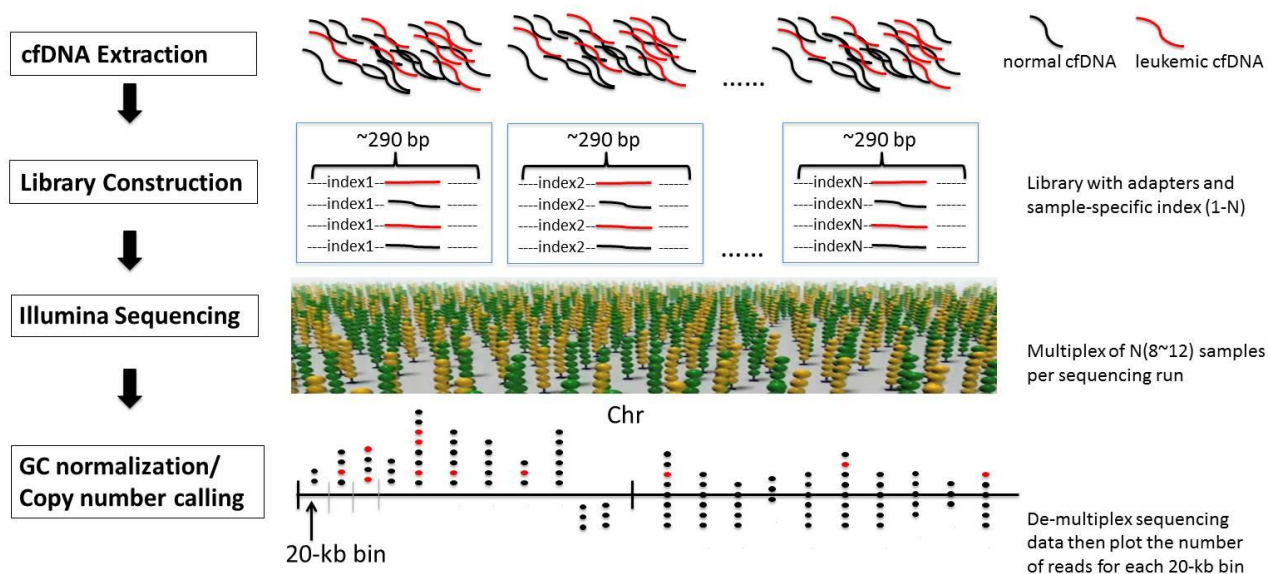


Figure 1. Workflow showing the four-step procedure for cfDSA: **1)** cfDNA extraction. The plasma cfDNA is a mixture of normal cfDNA (black curve lines) from healthy cells and leukemic cfDNA (red curve lines) from cancerous cells. **2)** Library construction. Each cfDNA sample is ligated with adapters and then amplified with one universal primer and one indexed primer. **3)** Illumina sequencing. 8-12 indexed samples are pooled together then sequenced in a HiSeq platform. **4)** Data analysis. Copy number callings are made through chromosome profiling.

Application of cfDSA on Ten Residual Samples

As illustrated in **Figure 1**, the workflow of cfDSA includes four steps, cfDNA extraction, library construction, Illumina platform HiSeq sequencing, and data analysis. For cfDNA extraction, residual leukemic blood (BL) and bone marrow (BM) samples were collected from ten cases (case #1, T-ALL-BL; #2, AML-BL; #3, CLL-BL; #4, MDS-BM; #5-6, MDS-BM; #7, myelofibrosis-BM; #8, hypereosinophilic syndrome-BL; #9, CLL-BL; and #10 AML-BL) following the recommended de-identification procedure.¹⁵ The collected samples were centrifuged at 1,600 g for 10 min, and 0.5-5 ml blood plasma or bone marrow fluid was moved into a new tube to re-centrifuge at 16,000 g for 10 min. The cfDNA was extracted using the circulating nucleic acid kit (Qiagen Inc., Valencia, CA) and further purified through

electrophoresis with 2% agarose gel followed by recovery using a gel purification kit (Qiagen Inc., Valencia, CA). The purified cfDNA was subject to library construction using DNA library preparation kit for Illumina (New England Biolabs, Inc, Ipswich, MA). Briefly, end repairing of cfDNA fragments was performed using T4 DNA polymerase and Klenow DNA polymerase. Then an adenine residue was added to the 3'-end of both strands by T4 polynucleotide kinase to form an overhang, which facilitated the ligation of adapters to the DNA fragments. The adapter-ligated DNA fragments were then enriched using a 15-cycle PCR with standard primers in the kit. For multiplexing purpose, a six-nucleotide index sequence was included in one primer. Each case had its own indexed primer. The resulted PCR product

was purified using Magnet beads (Beckman Coulter Inc., Pasadena, CA). The size distribution was determined using Bioanalyzer and quantified with real time qPCR (Qiagen Inc., Valencia, CA). The adapter-ligated DNA libraries were hybridized to the surface of sequencing flow cells. To multiplex, 8-12 samples in this study were mixed with other differently indexed samples. DNA clusters were generated using an Illumina cluster station, followed by 75 cycles of sequencing on an Illumina HiSeq 2500. The sequencing image processing and base calling were conducted following to Illumina's protocol by the Yale Center for Genome Analysis (YCGA).

The preprocessing, alignment, and chromosomal profiling were performed following Seqwise NGS Consulting Company's pipeline. Briefly, to obtain unique reads, the sequence data were aligned to NCBI36/hg18 assembly of Human Genome Browser. Standard post-sequencing alignment analysis was done using bowtie software to map the chromosomal locations of the sequenced cfDNA.¹⁶ If a

sequence read could be mapped to a unique location with two or less mismatches, it was determined to be a "unique" read. To translate the reads to a chromosomal profile, the reference genome excluding the repetitive regions was divided into about 125,000 20-kb-long bins. For each bin, the number of unique reads was determined. Because the GC content could affect the number of reads in each bin,⁶ the 125,000 bin reads were plotted against the GC content in each bin. The plot enables the calculation of the deviations from the GC content, and then the read numbers in each bin was normalized. The 125,000 normalized bin reads were plotted to a whole genome according to the bin location. For display purposes, the median of the 125,000 bin reads was set as 1 and all other bin read numbers were normalized to the median using robust Z score.^{17,18} The normalized bin reads were processed to generate a genome-wide chromosomal profile with normal two-copy regions in the median and abnormal copy number aberrations deviating from the median. Then the comparison of chromosome/FISH, aCGH, and the cfDSA results was conducted accordingly.

Table 1. The comparison of results from karyotyping/FISH, aCGH, and cfDSA on four abnormal cases.

| Case | Indication | Karyotype/FISH | aCGH (NCBI36/hg18, Mb) | cfDSA (NCBI36/hg18, Mb) |
|------|------------|--|---|---|
| 1 | T-cell ALL | 46,XY,dup(5)(p14.1p15.33), idic(6)(q12),idic(8)(p12), del(11)(q13.5q25), der(14)t(14;14)(q11;q32.1), add(19)(p13.2)[12/15] | deletions: 6p12p25.1(chr6:69.0-152.1), 8p23.3p12(chr8:0.2-34.1), 11q13.5q23.3(chr7:76.1-134.3). duplications: 5p15.33p14.1(chr5:2.0-26.9), 6p25.3q12(chr6:0.1-68.5), 8p12q24.3(chr8:34.2-146.2). | deletions: 6p12p25.1(chr6:68.9-152.1), 8p23.3p12(chr8:0.2-34.2), 11q13.5q23.3(chr7:76.3-134.4). duplications: 5p15.33p14.1(chr5:2.0-26.9), 6p25.3q12(chr6:0.1-68.9), 8p12q24.3(chr8:34.2-146.3). |
| 2 | AML | 46,XX,der(3)t(3;7)(q29;p14.1), del(5)(q23.1),-6,-7, +idic(8)(p21.1),+mar[14/15] | deletions: 5q23.1q35.3(chr5:119.0-180.6), 6p25.3p22.3(chr6:0.1-21.4), 6p12.3q27(chr6:46.1-170.7), 7p14.1q36.3(chr7:37.5-158.6). duplications: 6p22.2p21.1(chr6:24.9-44.5), 8p21.1q24.3(chr8:29.1-146.1). | deletions: 5q23.1q35.3(chr5:118.9-180.7), 6p25.3p22.3(chr6:0.2-24.5), 6p12.3q27(chr6:45.5-170.8), 7p14.1q36.3(chr7:37.5-158.8). duplications: 6p22.2p21.1(chr6:24.5-44.5), 8p21.1q24.3(chr8:28.0-146.3). |
| 3 | CLL | nuc ish(DLEU1x1,D13S1825x2, D12Z3x2)[52/200] | deletion: 13q14.2q21.1(chr13:47.4-53.9). | deletion: 13q14.2q21.1(chr13:46.2-54.1). |
| 4 | MDS | nuc ish(RELN,TES)x1[22/200], (MYCx2,PTPRTx1)[24/200] | Normal | Normal |

RESULTS

On average, two million raw reads were produced by Illumina HiSeq platform per sample, of which 1.5 million unique reads were sorted out by the bioinformatics processing. Parallel comparison of the results from karyotyping/FISH, aCGH, and cfDSA revealed that cfDSA results matched with diagnostic cytogenetic findings. Of the six samples (case #5-10) that have normal chromosome and FISH findings, no genomic aberrations were found in aCGH or cfDSA. The chromosome/FISH, aCGH and cfDSA results for four abnormal cases (#1-4) were summarized in **Table 1**. In two cases (#1 and 2) with clonal abnormalities featuring large chromosome deletions and duplications and in one case (#3) with a smaller chromosome deletion, aCGH and cfDSA detected similar genomic imbalances. For another case (#4) with a low percentage of abnormal cells, both aCGH and cfDSA failed to detect the abnormality.

The case 1 was a T-cell ALL patient, and its karyotype showed complex rearrangements of multiple chromosomes. Briefly, a stemline abnormal clone with additional materials onto a 5p and a 19p, a possible 6q/9p translocation, an isochromosome for the long arm of chromosome 8, an extra chromosome 11 with a deletion in its long arm, a loss of one chromosome 14, and a derivative chromosome 14 from a 14q11 and 14q32.1 translocation was noted in twelve cells and a sideline clone (subclone) with an additional aberration of an extra copy of chromosome 7 was seen in three cells. The aCGH revealed an abnormal pattern with a 26.7 Mb duplication of 5p15.33-5p14.1 (chr5:204,536-26,883,977), a 68.4 Mb duplication of 6p25.3-q12 (chr6:115,226-68,473,914), a 83.1 Mb deletion of 6q12-q25.1 (chr6:68,969,695-152,106,645), a 33.8 Mb deletion of 8p23.3-p12 (chr8:211,411-34,072,615), a 112.0 Mb

duplication of 8p12-q24.3 (chr8:34,289,110-146,250,965), and a 43.4 Mb deletion of 11q13.5-q25 (chr11:76,128,214-119,576,440). The karyotype was revised per aCGH findings: the additional material onto a 5p is most likely a complex segmental duplication and amplification of 5p14.1-p15.33; the possible 6q/9p translocation is likely an isodicentric

chromosome 6 with breakage and fusion at 6q12; the isochromosome 8 is an isodicentric chromosome 8 with breakage and fusion at 8p12; and the other chromosome aberrations may be balanced. All copy number aberrations detected by aCGH were also found in the cfDSA results (Figure 2).

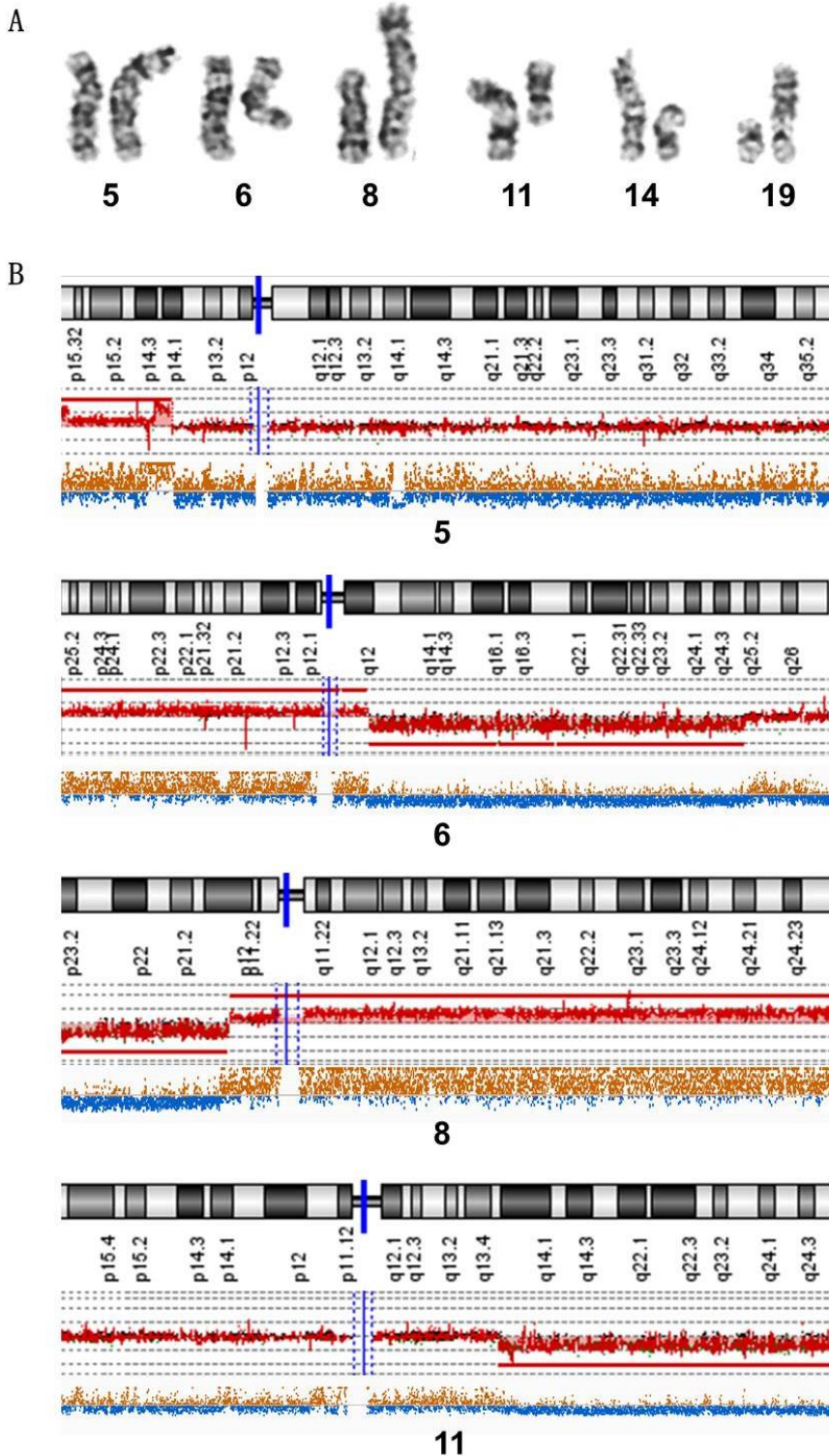


Figure 2. The comparison of results from karyotyping, aCGH and cfDSA on case 1. A) Clonal chromosome abnormality is detected. From left to right, 5p duplication/amplification, idic(6)(q12), idic(8)(p12), del(11)(q13.5q25), ?der(14), and ?add(19p). B) Comparison of the results between aCGH and cfDSA on the four abnormal chromosomes. For each chromosome, the top panel is an ideogram, the middle is a chromosome view from aCGH, and the lower is the chromosome profile from cfDSA.

The case 2 was an AML patient. Its karyotype revealed an abnormal clone featuring a 5q deletion, a translocation between chromosomal bands 3q29 and 7p14.1, losses of one copy of chromosomes 6 and 7, an extra copy of an isodicentric chromosome 8 with breakage and fusion at 8p21.1, and a marker chromosome. The aCGH analysis revealed an abnormal pattern with a 61.603 Mb deletion of 5q23.1-q35.3 (chr5:118,996,119-180,598,725), a 21.236 Mb deletion of 6p25.3-p22.3 (chr6:115,226-21,350,918), a 19.583 Mb duplication of 6p22.2-p21.1 (chr6:24,912,359-44,495,477), a 11.413 Mb deletion of 6p12.3-p11.2 (chr6:46,087,978-57,501,299), a 107.878 Mb deletion of 6q11.1-q27 (chr6:62,854,532-170,732,174), a 121.051 Mb deletion of 7p14.1-q36.3 (chr7:37,503,554-158,554,488), and a 116.945 Mb duplication of 8p21.1-q24.3 (chr8:29,147,311-146,092,035). Per aCGH results, the marker chromosome likely contains materials from 6p21.1-p22.2. All the abnormal regions detected with aCGH were also found in the cfDSA results (**Table 1**).

In case 3 with a FISH-detected 13q deletion present in 26% of blood leukocytes, aCGH detected a 6.6 Mb deletion of 13q14.2-q21.1 (chr13:47,357,404-53,920,330) and cfDSA detected a similar deletion (**Table 1**). For case 4 in which 10%~12% of bone marrow cells were detected by FISH to contain 7q and 20q deletions, both aCGH and cfDSA failed to detect copy number aberrations (**Table 1**).

DISCUSSION

The results from this pilot study demonstrated the technical and clinical feasibility of cfDSA. For three abnormal cases (#1-3), similar patterns of genomic copy number aberrations were detected from aCGH and cfDSA, and these genomic findings further confirmed and defined the clonal chromosome abnormalities. However, as shown in case #4, both aCGH on genomic DNA and cfDSA on bone marrow fluid cfDNA cannot detect low level of clonal abnormality.

Before the clinical application of cfDSA, several technical and clinical hurdles need to be overcome. First of all, there is a strong need for more in-depth investigation on the relationship between leukemic clonal abnormality in cellular level and leukemic portion of cfDNA from apoptotic leukemic cells. A better understanding of this relationship will be the biologic basis for its clinical application. Secondly, the analytic procedures need to be standardized and validated. The analytic resolution on the size of genomic copy number aberrations and on the percentage of leukemic portion of cfDNA detectable by next-generation sequencing needs to be established. For example, the cutoff of fetal cfDNA in maternal plasma was set at 4% for non-invasive prenatal screening for fetal aneuploidy (e.g. > 40 Mb for chromosome 21) using at least 10 million reads from next-generation sequencing.⁶⁻⁸ Similar to the validation of aCGH for clinical diagnosis,¹⁹ the analytical sensitivity and specificity could be defined using a receiver operating characteristic (ROC) statistics. Thirdly, clinical validity needs to be established using a large case series of each type of leukemia with normal and abnormal cytogenomic findings. As in the current cancer cytogenetic and aCGH

analyses, there is always a challenge to interpret transient chromosome abnormality²⁰ and define genomic content of complex chromosome findings such as bi-allelic aberrations, maker chromosome and double minutes.^{20,21} The introduction of cfDSA will reveal the missing part on tracking the disappearance or re-appearance of leukemic cells, but the interpretation of cfDSA results has to be based on cytogenomic findings. The cfDSA could be a part of an integrated analysis for cancer clonal evolution and disease progression.²² Lastly but not the least, there is no well-developed bioinformatics tools for cfDSA. In this pilot study, all data analyses depend on individual coding and testing which limits its use in a clinical setting. The development of a user-friendly designed and clinically oriented software for cfDSA will be a critical step to accelerate the data analysis and facilitate diagnostic interpretation.

In summary, this pilot study present a feasible method using next-generation sequencing on cfDNA to profiling chromosomal copy number aberrations for detecting leukemic clonal abnormalities. Albeit the caveat in the study due to limited number of cases and the low number of sequencing reads, the results have shown great potential for this new technology in detecting clonal abnormalities in neoplasm.

CONFLICT OF INTEREST

None.

ACKNOWLEDGEMENT

We thank Audrey Meusel who helped for proofreading our manuscript.

REFERENCES

- Mrozek K, Heerema NA, Bloomfield CD. Cytogenetics in acute leukemia. *Blood Rev.* 2004;18(2):115-136.
- Oscier DG. Cytogenetic and molecular abnormalities in chronic lymphocytic leukaemia. *Blood Rev.* 1994;8(2):88-97.
- Morris CM. Chronic myeloid leukemia: cytogenetic methods and applications for diagnosis and treatment. *Methods Mol Biol.* 2011;730:33-61.
- Sandberg AA, Meloni-Ehrigb AM. Cytogenetics and genetics of human cancer: methods and accomplishments. *Cancer Genet Cytogenet.* 2010;203(2):102-126.
- Bajaj R, Xu F, Xiang B, et al. Evidence-based genomic diagnosis characterized chromosomal and cryptic imbalances in 30 elderly patients with myelodysplastic syndrome and acute myeloid leukemia. *Mol Cytogenet.* 2011;4:3.
- Chiu RW, Chan A, Gao Y, et al. Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proc Natl Acad Sci USA.* 2008;105(51):20458-20463.
- Bianchi DW, Parker RL, Wentworth J, et al. DNA Sequencing versus standard prenatal aneuploidy screening. *N Engl J Med.* 2014;370(9):799-808.
- Xu Z, Xie J, Meng J, Li P, Pan X, Zhou Q. Non-invasive prenatal diagnosis: a comparison of cell free fetal DNA (cffDNA) based screening and fetal nucleated red blood cell (fnRBC) initiated testing. *N A J Med Sci.* 2013;6(4):194-199.
- Sozzi G, Conte D, Mariani L, et al. Analysis of circulating tumor DNA in plasma at diagnosis and during follow-up of lung cancer patients. *Cancer Res.* 2001;61(12):4675-4678.
- Ziegler A, Zangemeister-Wittke U, Stahel RA. Circulating DNA: a new diagnostic gold mine? *Cancer Treat Rev.* 2002;28(5):255-271.
- Gautschi O, Bigosch C, Huegli B, et al. Circulating deoxyribonucleic acid as prognostic marker in non-small-cell lung cancer patients undergoing chemotherapy. *J Clin Oncol.* 2004;22(20):4157-4164.
- Fleischhacker M, Schmidt B. Circulating nucleic acids (CNAs) and cancer—a survey. *Biochim Biophys Acta.* 2007;1775(1):181-232.

13. Schwarzenbach H, Hoon DS, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer*. 2011;11(6):426-437.
14. Gao Y, He Y, Yang Z, et al. Increased integrity of circulating cell-free DNA in plasma of patients with acute leukemia. *Clin Chem Lab Med*. 2010;48(1):1651-1656.
15. Grommisch B, DiAdamo A, Pan X, et al. Biobanking of residual specimens from diagnostic genetic laboratories: standard operating procedures, ethical and legal considerations, and research applications. *N A J Med Sci*. 2013;6(4):200-207.
16. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):R25.
17. Zhang JH, Chung TDY, Oldenburg KR. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J Biomol Screen*. 1999;4(2): 67-73.
18. Lai WR, Johnson MD, Kucherlapati R, Park PJ. Comparative analysis of algorithms for identifying amplifications and deletions in array CGH data. *Bioinformatics*. 2005;21(19):3763-3770.
19. Xiang B, Li A, Valentin D, Nowak NJ, Zhao H, Li P. Analytical and clinical validity of whole-genome oligonucleotide array comparative genomic hybridization for pediatric patients with mental retardation and developmental delay. *Am J Med Genet A*. 2008;146A(15):1942-1954.
20. Massaro S, Bajaj R, Pashankar F, et al. Bi-allelic deletions within 13q14 and transient trisomy 21 with absence of GATA1s in pediatric acute megakaryoblastic leukemia. *Pediatr Blood Cancer*. 2011;57(3):516-9
21. Kamath A, Tara H, Xiang B, Bajaj R, He W, Li P. Double minute MYC amplification and deletion of MTAP, CDKN2A, CDKN2B and ELAVL2 in an acute myeloid leukemia characterized by oligonucleotide-array comparative genomic hybridization. *Cancer Genet Cytogenet*. 2008;183(2):117-120.
22. Parisi F, Micsinai M, Strino F, et al. Integrated analysis of tumor samples sheds light on tumor heterogeneity. *Yale J Biol Med*. 2012;85(3):347-61.