Cancer Genetics 239 (2019) 22–25

Contents lists available at ScienceDirect

Cancer Genetics

journal homepage: www.elsevier.com/locate/cancergen

Short Communication

Rapid detection of chromosomal translocation and precise breakpoint characterization in acute myeloid leukemia by nanopore long-read sequencing

Chun Hang Au[a], Dona N. Ho[a], Beca B.K. Ip[a], Thomas S.K. Wan[b], Margaret H.L. Ng[b], Edmond K.W. Chiu[c], Tsun Leung Chan[a], Edmond S.K. Ma[a,*]

[a] Division of Molecular Pathology, Department of Pathology, Hong Kong Sanatorium and Hospital, 1/F Li Shu Fan Block, 2 Village Road, Happy Valley, Hong Kong
[b] Blood Cancer Cytogenetics and Genomics Laboratory, Department of Anatomical and Cellular Pathology, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong
[c] Honorary Consultant in Hematology and Hematological Oncology, Hong Kong Sanatorium and Hospital, Hong Kong

ARTICLE INFO

Article history:
Received 3 June 2019
Revised 21 July 2019
Accepted 21 August 2019

Keywords:
Acute myeloid leukemia
Cytogenetics and molecular genetics
Minimal residual disease
Nanopore sequencing
Structural variant

ABSTRACT

Detection of chromosomal translocation is a key component in diagnosis and management of acute myeloid leukemia (AML). Targeted RNA next-generation sequencing (NGS) is emerging as a powerful and clinically practical tool, but it depends on expression of RNA transcript from the underlying DNA translocation. Here, we show the clinical utility of nanopore long-read sequencing in rapidly detecting DNA translocation with exact breakpoints. In a newly diagnosed patient with AML, conventional karyotyping showed translocation t(10;12)(q22;p13) but RNA NGS detected *NUP98-NSD1* fusion transcripts from a known cryptic translocation t(5;11)(q35;p15). Rapid PCR-free nanopore whole-genome sequencing yielded a 26,194 bp sequencing read and revealed the t(10;12) breakpoint to be *DUSP13* and *GRIN2B* in head-to-head configuration. This translocation was then classified as a passenger structural variant. The sequencing also yielded a 20,709 bp sequencing read and revealed the t(5;11) breakpoint of the driver *NUP98-NSD1* fusion. The identified DNA breakpoints also served as markers for molecular monitoring, in addition to fusion transcript expression by digital PCR and sequence mutations by NGS. We illustrate that third-generation nanopore sequencing is a simple and low-cost workflow for DNA translocation detection.

Introduction

Detection of chromosomal translocation is clinically important for disease diagnosis, prognostication, therapeutic target identification, and monitoring of hematological malignancies. Conventional detection methods are cytogenetics, PCR and fluorescence *in situ* hybridization (FISH). Next-generation sequencing (NGS) is emerging as a powerful new tool in the detection of chromosomal translocations, such as massively parallel RNA sequencing to detect recurrent gene fusions in leukemia [1]. In theory, whole genome sequencing (WGS) by NGS offers an unbiased approach for the identification of rearrangements, but is hitherto considered impractical in clinical practice due to high cost and complexity. However third-generation sequencing platforms in particular the Oxford Nanopore MinION system allows the rapid sequencing of long reads and real-time base calling during sequencing at a reasonable cost. This method has been applied to detect structural variation by WGS in human genetic disorders [2]. More recently, nanopore sequencing has also been applied to the detection of leukemia-associated gene fusions coupled with anchored multiplex PCR for library preparation [3]. We report the clinical application of nanopore long-read sequencing for the rapid detection and precise breakpoint characterization of chromosomal translocations in acute myeloid leukemia (AML).

* Corresponding author.
E-mail addresses: tommy.ch.au@hksh.com (C.H. Au), dona.ny.ho@hksh.com (D.N. Ho), beca.bk.ip@hksh.com (B.B.K. Ip), thomaswan@cuhk.edu.hk (T.S.K. Wan), margaretng@cuhk.edu.hk (M.H.L. Ng), edmondkwchiu@gmail.com (E.K.W. Chiu), chris.tl.chan@hksh.com (T.L. Chan), eskma@hksh.com (E.S.K. Ma).

https://doi.org/10.1016/j.cancergen.2019.08.005

Materials and methods

Patient samples

Bone marrow samples were obtained from the patient at diagnosis (sample D in February 2017) and 6 follow-ups (samples F1 to F6, ranging from August 2017 to November 2018) (Supplementary Table S1). Peripheral blood sample was also obtained from the patient at diagnosis. DNA was extracted from each sample using QIAamp DNA Blood Mini Kit (Qiagen, Germany) or from derived buffy coat using EZ1 DNA Blood 350 µl Kit on EZ1 Advanced XL platform (Qiagen, Germany). Extracted DNA was quantified using Qubit dsDNA BR Assay Kit and Qubit 2.0 Fluorometer (Life Technologies, USA). RNA was extracted from each sample using QIAamp RNA Blood Mini Kit (Qiagen, Germany). Extracted RNA was quantified using Nanodrop 2000c Spectrophotometer (Thermo Fisher Scientific, USA). Pre-existing de-identified archival peripheral blood DNA and RNA samples were used as control in minimal residual disease monitoring study.

Cytogenetics and molecular cytogenetics

Conventional karyotyping of bone marrow sample (for leukemia karyotype) and peripheral blood sample (for constitutional karyotype) at diagnosis was performed as previously described [4]. Fluorescence *in situ* hybridization (FISH) was performed as previously described [4], using a *NUP98* dual-color break-apart probe (Empire Genomics).

Long-read nanopore sequencing

PCR-free whole-genome sequencing library was prepared using 500 ng DNA and a modified protocol of Rapid Sequencing Kit (SQK-RAD003, Oxford Nanopore Technologies, UK). Purification of 500 ng DNA was performed using 1x volume of AMPure XP beads (Beckman Coulter, USA) and elution from washed beads to 10 µL nuclease-free water at room temperature for 10 min. Purified DNA (436.5 ng in 7.5 µL) was subjected to transposome-based fragmentation and addition of adapters at 50 °C for 10 min. The library was sequenced using a MinION R9.4 flow cell (FLO-MIN106) for 44.5 h. Raw signals collected as FAST5 files were subjected to basecalling using Albacore version 2.1.3 (Oxford Nanopore Technologies, UK) on a Cray XC30 supercomputer. Sequencing reads were mapped to reference genome sequence hg19 using minimap2 [5] version 2.2-r424 (parameters `-ax map-ont`). Breakpoint-spanning sequencing reads across two different chromosomes of interest were identified by SAMtools [6] version 1.3.1.

Results

A 63-year-old Chinese woman presented with dyspnea on exertion and pallor. Complete blood counts showed: hemoglobin 10 g/dL, white cell count 76×10^9/L (53% blasts) and platelet count 157×10^9/L. Bone marrow examination showed 56% blasts and a morphological diagnosis of AML with maturation. The blast cells were positive for myeloperoxidase (MPO) and sudan black B on cytochemistry. Immunophenotyping showed expression of CD13, CD33, CD117, HLA-DR. and cytoplasmic MPO, but negative for B-cell and T-cell markers. Hence myeloid lineage was confirmed. Conventional molecular testing showed that the diagnostic bone marrow sample was positive for *FLT3* internal tandem duplication (ITD) (33 bp duplication) but negative for *FLT3* tyrosine kinase domain (TKD) mutation, *RUNX1-RUNX1T1* fusion transcript and *CBFB-MYH11* fusion transcript. Further DNA mutation screening by a 54-gene NGS myeloid panel according to a previously published protocol [7] confirmed *FLT3*-ITD of 33 bp in length at variant allelic

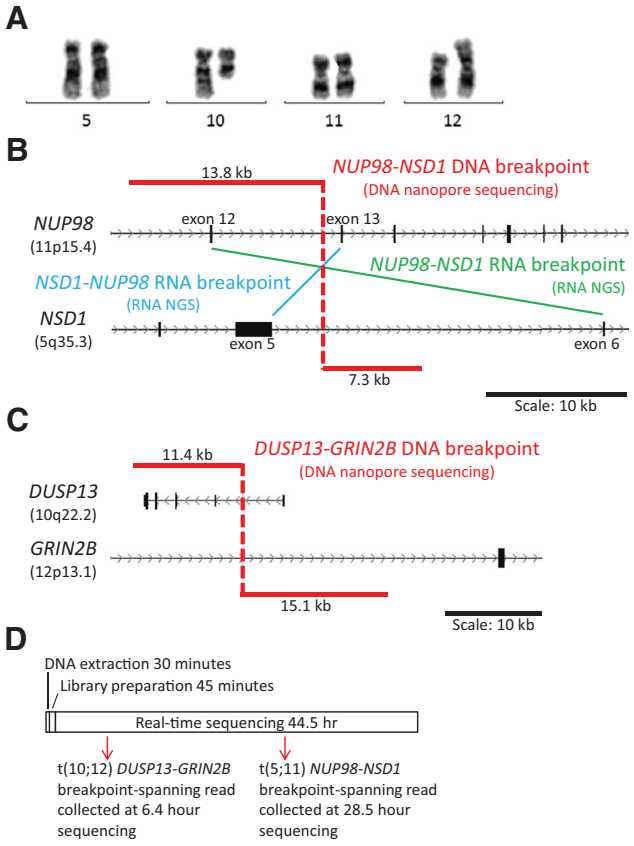

Fig. 1. Precise translocation breakpoint rapidly identified by nanopore long-read sequencing. (A) Partial karyotype of diagnostic bone marrow sample showed apparently normal chromosomes 5 and 11, and t(10;12)(q22;p13). (B) Genomic breakpoint of translocation between *NUP98* (chromosome 11) and *NSD1* (chromosome 5) was identified by a 20,709 bp sequencing read (red line), spanning 13.8 kb of *NUP98* and 7.3 kb of *NSD1*. The genomic breakpoint was consistent with *NUP98-NSD1* transcript breakpoint (green line) and *NSD1-NUP98* transcript breakpoint (blue line) identified by targeted RNA sequencing. (C) Genomic breakpoint of translocation between *DUSP13* (chromosome 10) and *GRIN2B* (chromosome 12) was identified by a 26,194 bp breakpoint-spanning read. The head-to-head gene fusion was predicted to be non-productive in terms of RNA transcription and indirectly supported by absence of *GRIN2B*-involving fusion transcript (targeted by RNA sequencing). Gene transcription direction was indicated as arrows within gene structure diagram. (D) Timeline of experimental workflow from bone marrow sample to genomic breakpoint identification.

fraction (VAF) of 29.2% but did not detect other mutations. Cytogenetics of bone marrow sample (Fig. 1A) showed 46,XX,inv(9)(p11q13)c,t(10;12)(q22;p13)[17]/47,XX,+8,inv(9)(p11q13)c[2]/46,XX,inv(9)(p11q13)c[3]. The t(10;12)(q22;p13) and trisomy 8 occurred in separate clones. Constitutional nature of the t(10;12)(q22;p13) translocation was excluded by phytohemagglutinin-stimulated peripheral blood lymphocyte culture, which showed normal karyotype and constitutional inv(9)(p11q13) only.

To detect potential driver gene fusion event that may arise from t(10;12)(q22;p13), targeted RNA NGS panel of 1385 genes was performed on the diagnostic bone marrow sample in accordance with a previously published protocol [4]. However, no fusion transcript involving the visualized translocation breakpoints was identified. Unexpectedly, *NUP98-NSD1* leukemic fusion transcript and *NSD1-NUP98* reciprocal fusion transcript were identified instead (Fig. 1B), based on 103 bp and 136 bp breakpoint contig sequences reconstructed by 10 and 26 reads, respectively. Breakpoint sequence analysis showed fusion of *NUP98* exon 12 to *NSD1* exon 6 for the *NUP98-NSD1* transcript and fusion of *NSD1* exon 5 to *NUP98* exon 13 for the reciprocal *NSD1-NUP98* transcript, which were the

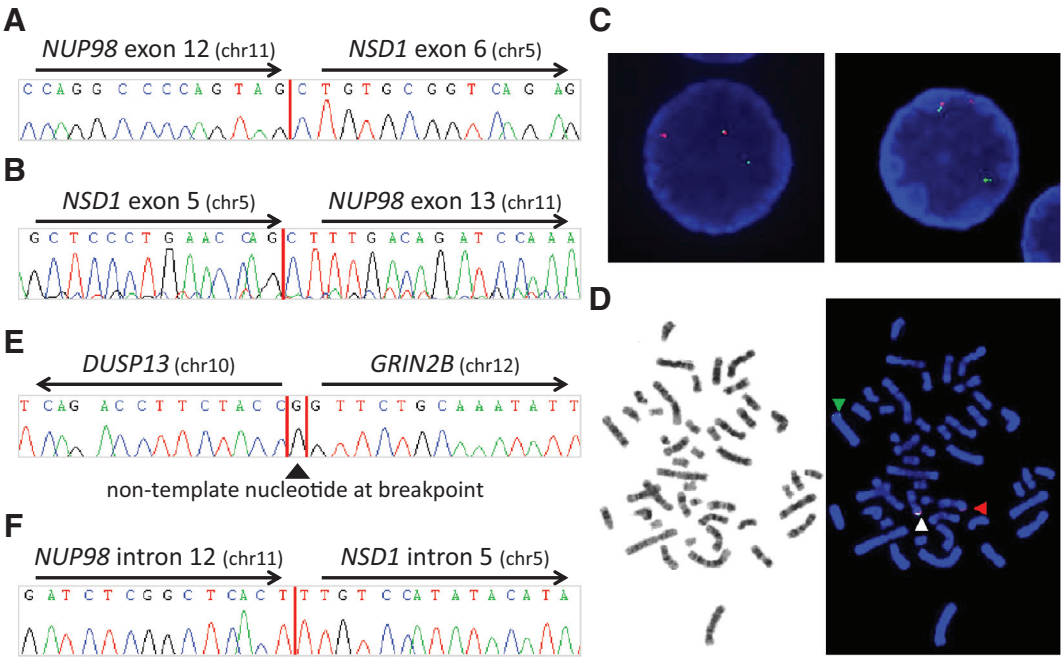

Fig. 2. Orthogonal confirmation of chromosome translocations detected by next-generation sequencing (NGS) and nanopore sequencing. (A) Sanger sequencing chromatogram of *NUP98-NSD1* RT-PCR product confirmed the fusion breakpoint detected by RNA NGS (red line). Gene transcription direction was indicated by black arrows. (B) Sanger sequencing chromatogram of *NSD1-NUP98* RT-PCR product confirmed the fusion breakpoint detected by RNA NGS (red line). Gene transcription direction was indicated by black arrows. (C) Interphase fluorescence *in situ* hybridization (FISH) on cell nuclei showed *NUP98* gene rearrangement. A *NUP98* dual-color break-apart probe was used (chromosome 11p15.4). Split red (*NUP98* 5′ region) and green (*NUP98* 3′ region) signals indicated gene rearrangement while fusion signals indicated an intact gene. (D) Metaphase FISH with the same probe and correlation with G-banding showed intact fusion signal on 11p15 (white arrowhead), split red signal on derivative chromosome 11p (red arrowhead) and split green signal on derivative chromosome 5q (green arrowhead). (E) Sanger sequencing chromatogram of breakpoint PCR product confirmed the translocation breakpoint (red lines) detected by nanopore sequencing. The breakpoint on derivate chromosome 10 was seq[GRCh37] t(10;12)(q22.2;p13.1) g.[chr10:pter_cen_76864414::G::chr12:pter_13933686inv]. (F) Sanger sequencing chromatogram of breakpoint PCR product confirmed the translocation breakpoint (red line) detected by nanopore sequencing. The breakpoint on derivate chromosome 11 was seq[GRCh37] t(5;11)(q35.3;p15.4) g.[chr5:176642857_qterinv::chr11:3757799_cen_qter]. Gene transcription direction was indicated by gray arrows.

same RNA breakpoints previously described in the literature [8]. These NGS results were confirmed by RT-PCR and Sanger sequencing (Fig. 2A and B). Detection of the two fusion transcripts inferred a balanced translocation between chromosomes 5 and 11 in the form of t(5;11)(q35;p15), which was a known cryptic translocation not amendable for detection by routine karyotyping. Translocation involving *NUP98* was further supported by both interphase and metaphase FISH by a *NUP98* dual-color break-apart probe (Fig. 2C and D).

Although a driver gene fusion event *NUP98-NSD1* was identified, the nature of t(10;12)(q22;p13) remained elusive. Long-read PCR-free WGS of bone marrow sample was performed on a MinION sequencer to characterize the translocation at DNA level. At 6.4 h of the real-time sequencing run, a 26,194 bp sequencing read was collected (Fig. 1D) and revealed the exact translocation breakpoint between *DUSP13* (chromosome 10q22.2) and *GRIN2B* (chromosome 12p13.1) at single-nucleotide resolution (Fig. 1C). More importantly, both loci were split at the intron regions and joined together in head-to-head configuration (transcriptional direction of annotated RefSeq transcripts). It was predicted that this translocation was non-productive in terms of RNA transcription. Correlation of breakpoint genes with the targeted RNA NGS panel design showed that *GRIN2B* was targeted, but the absence of detectable fusion transcript indirectly supported the non-productive nature of the translocation. The translocation breakpoint on the detected derivative chromosome 10 revealed by long-read sequencing was confirmed by conventional PCR and Sanger sequencing to be t(10;12)(q22.2;p13.1) g.[chr10:pter_cen_76864414::G::chr12:pter_13933686inv] (Fig. 2E). The *GRIN2B* gene encodes a glutamate-activated ion channel found at excitatory synapses throughout the brain, while the *DUSP13* gene encodes two distinct dual-specificity phosphatases (DSP), testis and skeletal muscle-specific (DSP) and muscle-restricted DSP. Due to non-productive nature of the gene fusion and no direct involvement of the gene products in normal hemopoiesis, it was concluded that t(10;12)(q22;p13) was a passenger structural variant (SV), although altered transcriptional activity due to translocation-mediated gene disruption is not entirely excluded.

As the real-time long-read sequencing progressed, the exact translocation breakpoint between *NUP98* (chromosome 11p15.4) and *NSD1* (chromosome 5q35.3) was also identified by a 20,709 bp sequencing read collected at 28.5 h of the sequencing run (Fig. 1D). *NUP98* intron 12 was fused to *NSD1* intron 5 in tail-to-head direction (Fig. 1B). The DNA-level translocation breakpoint and configuration revealed by long-read DNA sequencing was concordant with the RNA-level information revealed by RNA NGS. The DNA translocation breakpoint on the detected derivative chromosome 11 was confirmed by conventional PCR and Sanger sequencing to be t(5;11)(q35.3;p15.4) g.[chr5:176642857_qterinv::chr11:3757799_cen_qter] (Fig. 2F). It was concluded that translocation t(5;11)(q35;p15) was indeed present and considered as a driver SV. The *FLT3*-ITD originally detected by conventional molecular testing and NGS was also detected by a 20,363 bp nanopore sequencing read.

The patient was treated by 7:3 and reassessment showed non-remission, but achieved complete remission (CR) after changing to sorafenib and homoharringtonine [9]. Molecular monitoring showed that the *FLT3*-ITD was suppressed by sorafenib (Supplementary Fig. S1). The patient was maintained on sorafenib and in CR for 13 months, after which the disease relapsed. Repeat NGS

testing showed resurgence of *FLT3*-ITD and appearance of *FLT3*-TKD D835V mutation at VAF of 23%, in addition to *NRAS* and different *WT1* mutations. The *FLT3*-TKD was not found at diagnosis but detected at relapse on sorafenib. This phenomenon was most probably due to lack of inhibitory effect of type II FLT3-inhibitor such as sorafenib to *FLT3*-TKD, in contrast to type I FLT3-inhibitor such as midostaurin that could inhibit both ITD and TKD [10]. The patient was treated by low-dose ara-C, VP16, midostaurin and venetoclax, but the leukemia was persistent and she succumbed 21 months after initial diagnosis.

Minimal residual disease (MRD) monitoring of 6 follow-up diagnostic bone marrow samples was performed serially (Supplementary Fig. S1, Supplementary Fig. S2 and Supplementary Fig. S3). The same 54-gene NGS myeloid panel was applied to the follow-up samples to monitor VAF of DNA mutations (mutation detected at diagnosis and any new mutations) (Supplementary Table S1). Although the *FLT3*-ITD was suppressed by sorafenib (Supplementary Fig. S1), both t(5;11) derivative chromosome 11 and t(10;12) derivative chromosome 10 were persistently detectable by specific breakpoint PCR (Supplementary Fig. S2). Also *NUP98-NSD1* transcript droplet digital PCR showed persistent detection at level comparable to the diagnostic sample (Supplementary Fig. S3, Supplementary Fig. S4 and Supplementary Table S2). The eventual hematological relapse occurred in parallel with the molecular relapse as detected by the NGS myeloid panel. Taken together, the MRD results of our patient showed that undetectable *FLT3*-ITD status was not predictive of long-term disease remission and confirmed the recommendation that it should not be used alone as a single marker of MRD [11]. The comprehensive characterization of SV by the NGS approach in our patient provided useful MRD marker for molecular monitoring.

Discussion

Although NGS is able to detect short variations, namely single-nucleotide variants, insertions and deletions at single-base resolution, the short read length (<1 kb) is limiting its clinical application for SV detection. This limitation is partially circumvented by advanced NGS library preparation techniques, such as targeted RNA sequencing, mate-pair reads of circularized fragments, microfluidics-partitioned linked reads.

Nanopore sequencing, by delivering unprecedentedly long read length (1–1000 kb), is an emerging third-generation sequencing technology [12]. Recent technology improvement pushes nanopore sequencing capacity to gigabase level and renders the technology appealing for SV detection. Even though nanopore sequencing accuracy at the single-nucleotide level is currently lower than short-read NGS, SV detection from the long reads is robust as demonstrated by the orthogonal validation in this and other studies [13].

There is also immediate availability of data for sequencing, since real-time base calling is practiced for resulting reads to be mapped and analyzed while the run is on-going [14]. In this study, the key sequencing reads spanning two translocations were detected at 6.4 and 28.5 h of sequencing. Since the nanopores randomly sequence DNA molecules from the sample, the time of collecting such sequences in other sequencing run is expected to be different (any time from the beginning to 44.5 h in this study). Such timing and sequencing depth depends on the SV abundance in the sample, sample DNA quality and actual sequencing throughput obtained. Systematic correlation of these factors is important in clinical setting and requires further study. After detecting known SV of interest with precise breakpoint, orthogonal confirmation was needed and performed (Fig. 2). The nanopore sequencing *per se* is hitherto not for general SV screening without further validation study.

This platform is particularly disruptive in clinical setting because of negligible capital cost, minimal instrumentation, simple library preparation and relatively routine bioinformatics workflow. In terms of total cost from DNA sample to sequencing reads, the negligible capital cost of nanopore sequencing is in marked contrast to that of mainstream NGS platforms. Library preparation and sequencing reagent cost for a single-sample WGS by nanopore sequencing is also lower than NGS. Therefore, the nanopore sequencing workflow can be readily reproduced in other molecular pathology laboratories with minimal effort. It complements other existing techniques for comprehensive molecular profiling of myeloid neoplasms (Supplementary Table S3). Our patient case shows that nanopore sequencing allows rapid detection of chromosomal translocation and precise breakpoint characterization, which benefits clinical management.

Declaration of Competing Interest

None.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cancergen.2019.08.005.

References

[1] Kim B, Lee H, Shin S, Lee ST, Choi JR. Clinical evaluation of massively parallel RNA sequencing for detecting recurrent gene fusions in hematologic malignancies. J Mol Diagn 2019;21(1):163–70.
[2] Cretu Stancu M, van Roosmalen MJ, Renkens I, Nieboer MM, Middelkamp S, de Ligt J, et al. Mapping and phasing of structural variation in patient genomes using nanopore sequencing. Nat Commun 2017;8:1326.
[3] Jeck WR, Lee J, Robinson H, Le LP, Iafrate AJ, Nardi V. A nanopore sequencing-based assay for rapid detection of gene fusions. J Mol Diagn 2019;21(1):58–69.
[4] Ma ESK, Wan TSK, Au CH, Ho DN, Ma SY, Ng MHL, et al. Next-generation sequencing and molecular cytogenetic characterization of ETV6-LYN fusion due to chromosomes 1, 8 and 12 rearrangement in acute myeloid leukemia. Cancer Genet 2017;218–219:15–19.
[5] Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 2018;34(18):3094–100.
[6] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics 2009;25(16):2078–9.
[7] Au CH, Wa A, Ho DN, Chan TL, Ma ES. Clinical evaluation of panel testing by next-generation sequencing (NGS) for gene mutations in myeloid neoplasms. Diagn Pathol 2016;11(1):11.
[8] Shiba N, Ichikawa H, Taki T, Park M, Jo A, Mitani S, et al. *NUP98-NSD1* gene fusion and its related gene expression signature are strongly associated with a poor prognosis in pediatric acute myeloid leukemia. Genes Chromosomes Cancer 2013;52:683–93.
[9] Lam SS, Ho ES, He BL, Wong WW, Cher CY, Ng NK, et al. Homoharringtonine (omacetaxine mepesuccinate) as an adjunct for FLT3-ITD acute myeloid leukemia. Sci Transl Med 2016;8(359):359ra129.
[10] Larrosa-Garcia M, Baer MR. FLT3 inhibitors in acute myeloid leukemia: current status and future directions. Mol Cancer Ther 2017;16(6):991–1001.
[11] Schuurhuis GJ, Heuser M, Freeman S, Béné MC, Buccisano F, Cloos J, et al. Minimal/measurable residual disease in AML: a consensus document from the European Leukemianet MRD Working Party. Blood 2018;131:1275–91.
[12] Loose MW. The potential impact of nanopore sequencing on human genetics. Hum Mol Genet 2017;26(R2):R202–7.
[13] Jain M, Koren S, Miga KH, Quick J, Rand AC, Sasani TA, Tyson JR, Beggs AD, Dilthey AT, Fiddes IT, Malla S, Marriott H, Nieto T, O'Grady J, Olsen HE, Pedersen BS, Rhie A, Richardson H, Quinlan AR, Snutch TP, Tee L, Paten B, Phillippy AM, Simpson JT, Loman NJ, Loose M. Nanopore sequencing and assembly of a human genome with ultra-long reads. Nat Biotechnol 2018;36:338–45.
[14] Loose M, Malla S, Stout M. Real-time selective sequencing using nanopore technology. Nat Methods 2016;13(9):751–4.