Panel testing for gene mutations in acute myeloid leukaemia by next-generation sequencing

Chun Hang Au*, Wing Yi Wa*, Dona N Ho, Tsun Leung Chan, Edmond S K Ma#
Department of Pathology, Hong Kong Sanatorium & Hospital, Happy Valley, Hong Kong
*Equal contribution #Correspondence to Dr Edmond S K Ma (eskma@hksh.com)

Introduction

Genomic techniques in recent years have allowed the identification of many mutated genes that are important in the pathogenesis of acute myeloid leukaemia (AML). Together with cytogenetics aberrations, these gene mutations are powerful prognostic markers in AML and can be used to select patients for optimal post-remission therapy. Emerging data also show that these mutations may serve as predictive markers of treatment dose or choice upfront. The mutated genes hold promise as therapeutic targets themselves.

We aim to evaluate the applicability of a gene panel for the detection of AML mutations in a diagnostic molecular pathology laboratory.

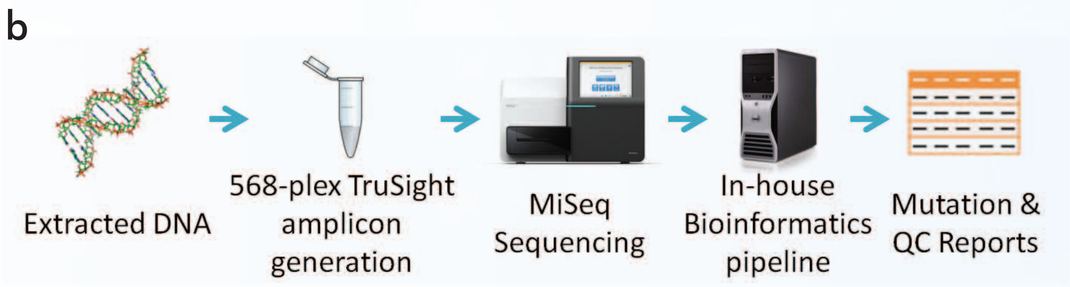
Materials & Methods

A total of 31 AML samples (bone marrow = 25 and peripheral blood = 6) were accrued for the study. They comprised 11 males and 20 females at a median age of 58 years. The diagnoses were M0 = 1, M1 = 6, M2 = 7, M3 = 1, M4 = 1, M5 = 2, M6 = 2, NOS = 2, AML with MDS related changes = 2, AML transformed from MDS or MDS/MPD = 4, therapy-related AML = 1 and refractory AML = 2. DNA was extracted from the corresponding specimens. Germline DNA from buccal swab was available in one patient. A total of 54 genes (full coding exons of 15 genes and exonic hotspots of 39 genes) were targeted by 568 amplicons (length range: 225 – 275bp) (Fig. **1a**). The combined coverage was 141kb. Amplicon sequencing libraries were prepared from 50 ng of DNA per sample using TruSight myeloid sequencing panel (Illumina, USA). Paired-end sequencing runs were performed on a MiSeq (Illumina, USA) with reagent kit v3. Sequences obtained were analyzed by in-house bioinformatics pipeline, which mainly consisted of BWA, SAMtools, Ensembl Variant Effect Predictor and a novel algorithm to detect long insertion in amplicon sequencing data. No confirmation by Sanger sequencing was performed (Fig. **1b**).

a

Gene	Target Region (exon)	Gene	Target Region (exon)	Gene	Target Region (exon)	Gene	Target Region (exon)
ABL1	4–6	*DNMT3A*	full	*KDM6A*	full	*RAD21*	full
ASXL1	12	*ETV6/TEL*	full	*KIT*	2, 8–11, 13 + 17	*RUNX1*	full
ATRX	8–10 and 17–31	*EZH2*	full	*KRAS*	2 + 3	*SETBP1*	4 (partial)
BCOR	full	*FBXW7*	9 + 10 + 11	*MLL*	5–8	*SF3B1*	13–16
BCORL1	full	*FLT3*	14 + 15 + 20	*MPL*	10	*SMC1A*	2, 11, 16 + 17
BRAF	15	*GATA1*	2	*MYD88*	3–5	*SMC3*	10, 13, 19, 23, 25 + 28
CALR	9	*GATA2*	2–6	*NOTCH1*	26–28 + 34	*SRSF2*	1
CBL	8 + 9	*GNAS*	8 + 9	*NPM1*	12	*STAG2*	full
CBLB	9, 10	*HRAS*	2 + 3	*NRAS*	2 + 3	*TET2*	3–11
CBLC	9, 10	*IDH1*	4	*PDGFRA*	12, 14, 18	*TP53*	2–11
CDKN2A	full	*IDH2*	4	*PHF6*	full	*U2AF1*	2 + 6
CEBPA	full	*IKZF1*	full	*PTEN*	5 + 7	*WT1*	7 + 9
CSF3R	14–17	*JAK2*	12 + 14	*PTPN11*	3 + 13	*ZRSR2*	full
CUX1	full	*JAK3*	13				

b



c

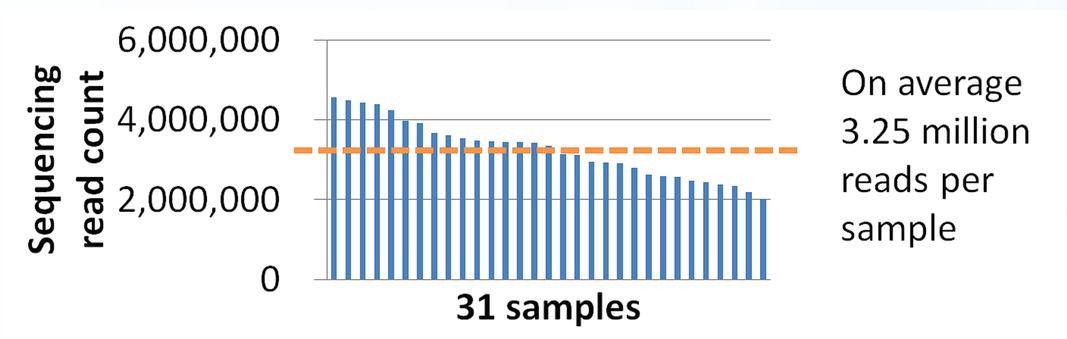


Figure 1
AML gene mutation panel testing. **(a)** Full exons of 15 genes and exonic hotspots of 39 genes are targeted for sequencing. Mutated genes among the samples (green) and in-house mutation positive control samples (based on conventional methods; orange) are highlighted. **(b)** Overview of panel testing workflow. **(c)** Uniformity of sequencing reads obtained per sample.

Results

The mean count of sequencing reads obtained per sample was 3.25 million (range 2.02 – 4.57 million) and the mean sequencing depth was over 3000X **(Fig. 1c)**. Thirty-eight mutations in 17 genes were detected in 20 of 31 samples (65%) **(Fig. 2)**. On average 1.9 mutations (range 1 – 5) were detected per sample. Mutations were detected in the following genes: *FLT3* (n = 7), *NPM1* (n = 7), *IDH2* (n = 3), *RAD21*, *CSF3R*, *STAG2*, *DNMT3A*, *TET2*, *IDH1*, *KIT* (all n = 2), *NRAS*, *ETV6/TEL*, *CEBPA*, *IKZF1*, *TP53*, *SF3B1 and SMC3* (all n = 1). Mutations detected by NGS were compared with previous results from conventional methods if available. NGS results were not just concordant with conventional results but also more sensitive in detecting additional mutations at low allelic burden. Detection of long insertion mutations (e.g. >70bp *FLT3* internal tandem duplication in the short amplicon sequencing data was challenging and required new bioinformatics algorithm **(Fig. 3)**.

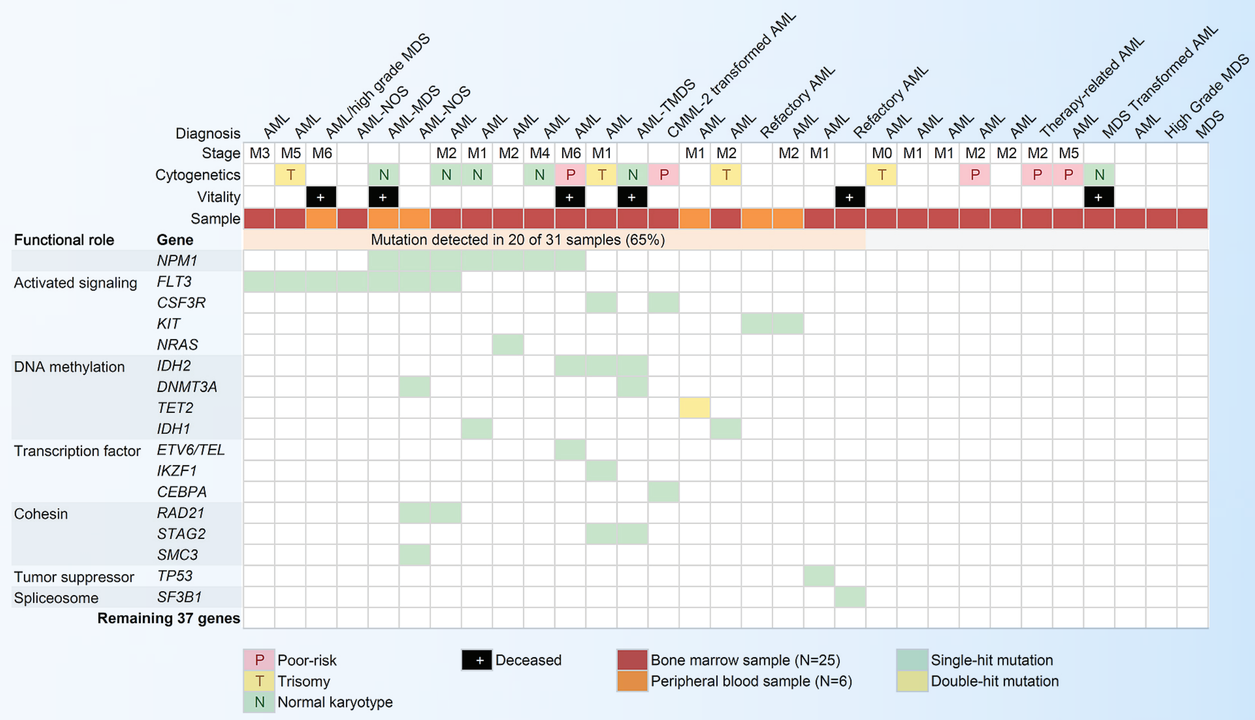


Figure 2
Mutation status matrix across 31 samples and 54 genes, arranged in columns and rows, respectively. Types of cytogenetics results (P: poor-risk; T: trisomy; N: normal karyotype), vitality (+: deceased), sample (red: bone marrow; orange: peripheral blood) and mutation (green: single-hit; yellow: double-hit) are represented by colored boxes.

Figure 3
Challenges in detecting *FLT3* internal tandem duplication (ITD) in amplicon sequencing reads. **(a)** In case of short ITD, unaligned duplicated sequence is located in the middle of sequencing reads. Insertion variation can be readily called by mainstream variation callers. **(b)** In case of long ITD (>70bp), unaligned duplicated sequence is too long to be fully sequenced and is usually soft-clipped and ignored by variation callers. A new bioinformatics algorithm was developed to extract the unaligned sequence for realignment so that the long ITD can be detected.

Conclusion

Comprehensive 54-gene panel testing revealed a high frequency and a diverse spectrum of mutations in AML. Gene panel testing by NGS approach in a diagnostic molecular pathology laboratory allows timely, sensitive and accurate detection of actionable AML gene mutations to individualize patient management.