Detection of internal tandem duplications in FLT3 with Anchored Multiplex PCR and next-generation sequencing

Introduction

FLT3 encodes a receptor tyrosine kinase that is involved in p53 activation and has roles in cell growth arrest and apoptosis. Internal tandem duplications (ITDs) in the juxtamembrane domain result in constitutive activation of FLT3, causing aberrant cell growth leading to tumorigenesis.[1] *FLT3*-ITDs are associated with poor prognosis of acute myeloid leukemia (AML) and are detected in about 25% of AML cases.[1-3] As *FLT3*-ITD expressed kinases are sensitive to tyrosine kinase inhibitors, they are of considerable interest for the development of novel AML treatments.[4,5]

ITDs arise in DNA from duplication of a region that is inserted adjacent to its origin. Sometimes, an intervening sequence is inserted between the original and duplicated segments. In *FLT3*, ITDs typically arise in exon 14 and can span intron 14 and exon 15.[6] The standard method to detect ITDs is capillary gel electrophoresis (CGE), which detects the size difference in a *FLT3* PCR amplicon that arises due to the presence of an ITD. Although CGE is regarded as an accurate and reliable method to detect ITDs, this method cannot be easily coupled with assays that detect other mutation types common in AML. This is important because 14-25% of *FLT3*-ITD-positive AML cases have 2 or more mutations in the *FLT3* gene, which is believed to contribute to the variable success rates of FLT3 inhibitors.[7,8] In addition, a recent study showed that AML cases contain a median of 4 mutations occurring across multiple gene targets, with *FLT3*, *NPM1, CEBPA, TP53, DNMT3A* and *RUNX1* identified as particularly important predictors of outcomes.[3] For example, NPM1 mutations that co-exist with *FLT3*-ITDs are associated with less favorable outcomes than without *FLT3*-ITDs. Therefore, simultaneous detection of multiple types of mutations across multiple target genes would be the most efficient and economical approach to characterize AML samples.

Next-generation sequencing (NGS) enables comprehensive detection of multiple mutation types. However, detection of ITDs by NGS is particularly challenging for multiple reasons. First, since ITD and reference sequences are identical, sequenced fragments that do not contain flanking regions cannot be distinguished from one another. Second, the sequence intervening the ITD and duplicated region may not have sequence similarity to FLT3 or may be too short to map, making alignment of fragments to the reference sequence challenging. Finally, ITDs are very diverse, ranging from ~3–400bp in length and inserting at multiple chromosomal positions.[9,10] This renders opposing primer-based NGS enrichment methods difficult, since primers must be designed to amplify the ITD-containing region without prior knowledge of its size and insertion position. These approaches are further limited by the read length required to amplify longer ITDs. Alternatively, hybrid capture-based methods cannot distinguish between original and duplicated sequences unless the sequenced read overlaps the breakpoint, which is not known prior to interrogation.

Results

We developed VariantPlex® Myeloid targeted NGS assays to detect *FLT3*-ITDs, in addition to other variants, from clinical-type genomic DNA samples. These assays are based on Anchored Multiplex PCR (AMP™), a target enrichment strategy for NGS that uses molecular barcoded (MBC) adapters and single gene-specific primers (GSPs) for amplification, permitting open-ended capture of DNA fragments from a single end. We designed GSPs to amplify the commonly mutated juxtamembrane domain and tyrosine kinase domain 1 (TKD1) of *FLT3*. Because the MBC adapters contain universal primer binding sites, amplification from GSPs is unrestricted by opposing primers and can amplify both large and small fragments, without prior knowledge of downstream sequences. This ability to amplify small, degraded fragments of DNA increases read depth and coverage of target regions from low-input samples. Importantly, since AMP primers function independently of each other, multiple overlapping "snapshots" of the ITD-containing region are obtained from both directions. As such, anchored reads originating from redundant, yet independent, GSPs contained in the VariantPlex Myeloid panels enable complete coverage and de novo assembly of variable ITD-containing regions (**Fig. 1** on the following page).

Fig. 1A

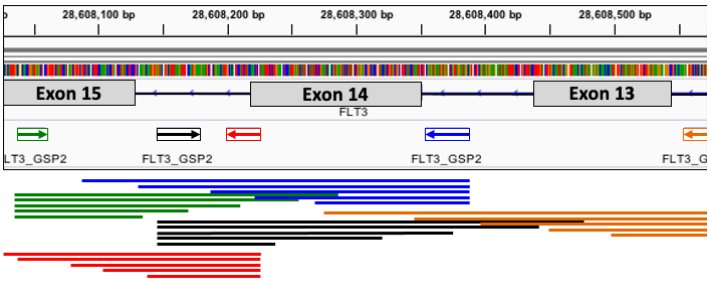

Fig. 1B

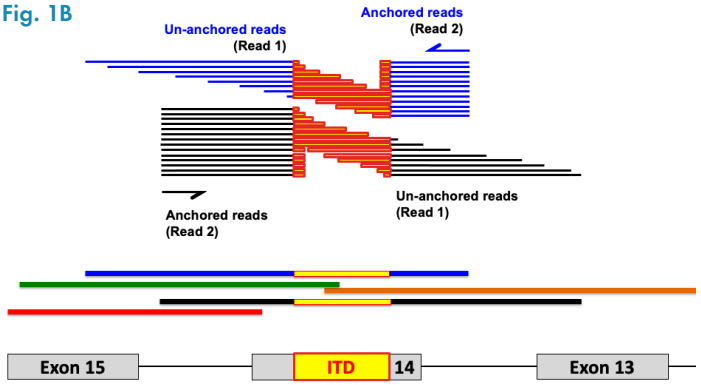

Figure 1. Open-ended amplification from independent gene-specific primers enables de novo assembly of ITD- containing regions. **(1A)** Cartoon depiction of anchored reads originating from GSPs contained in the VariantPlex Myeloid panels (not all GSPs are shown). **(1B)** Example illustration of de novo sequence assembly of ITD-containing region, with ITD shown in yellow.

Fig. 1A Anchored reads combined with primer redundancy enable:

- Complete coverage of ITD region and all ITDs of any size and insertion point
- Robust and accurate de novo assembly

Fig. 1B

- Reads are clustered by originating primer
- Overlapping read mate pairs are merged
- Consensus is generated based on sequence similarity
- Consensus reads from the same gene are merged to generate ITD content, length and insertion point

Fig. 2A

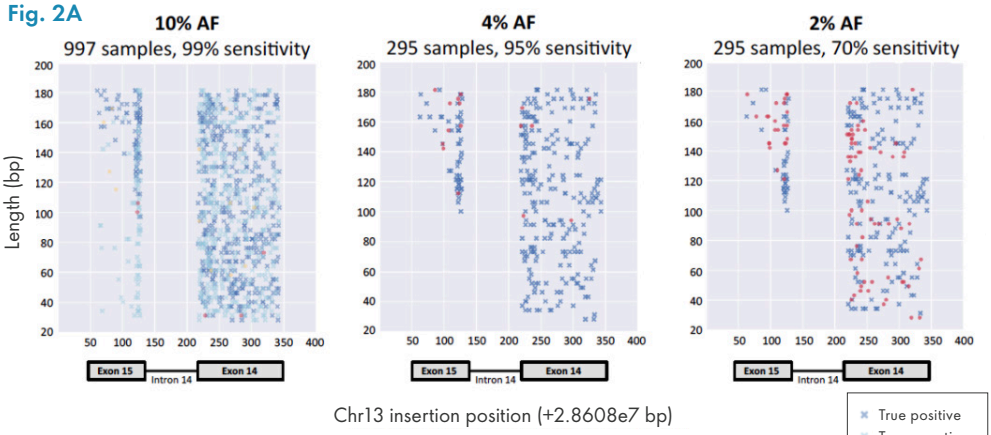

Fig. 2B

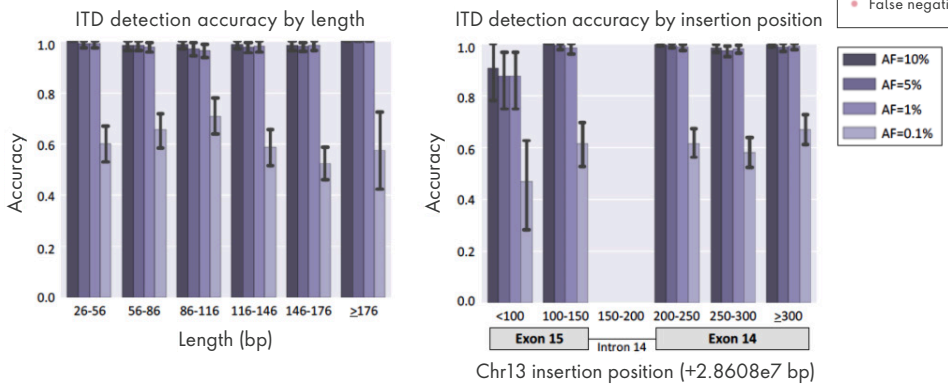

Figure 2. Accuracy of FLT3-ITD detection by length, chromosome 13 insertion position, and allele frequency (AF). In silico datasets were generated to represent the spectrum of known ITDs in FLT3 with varying lengths and insertion positions in chromosome 13, as well as varying AFs. Each in silico dataset was constructed to simulate reads originating from GSPs in the VariantPlex Myeloid panels. In silico datasets were analyzed with the Archer Analysis variant pipeline, and each analyzed sample was identified as true positive, false positive, true negative or false negative.

(2A) Scatter plots of detected (true and false positive) and undetected (true and false negative) ITDs with the indicated lengths and insertion positions are shown for 10%, 4% and 2% AFs. 97.4% specificity was observed at 10% AF and would be expected to increase with decreasing AF, and therefore only sensitivity data (true positive and false negative calls) are shown for the plots representing 2% and 4% AFs.

(2B) Accuracy of ITD detection as a function of length (left) or insertion position (right) from 0.1% to 10% AF.

To further test our detection algorithm on the vast space of possible *FLT3*-ITDs, we generated over 2000 in silico datasets containing *FLT3*-ITDs of a variety of lengths, insertion points, and allele fractions. These in silico datasets were used to refine and, ultimately, measure the sensitivity and specificity of our novel de novo assembly algorithm. Each in silico dataset was constructed to simulate reads originating from GSPs in the VariantPlex Myeloid panels. In silico datasets were analyzed with the Archer® Analysis variant pipeline, and each ITD was identified as true positive, false positive, true negative or false negative (**Fig. 2**). These data show that the Archer Analysis variant pipeline detects ITDs

with near 100% specificity (>97.4%) for all AFs tested, and detects ITDs with >95% sensitivity down to 4% AF, and 70% sensitivity at 2% AF (**Fig. 2A**). This results in near 100% detection accuracy of ITDs as low as 1% AF regardless of ITD length and across most insertion positions (a slight decrease in accuracy is observed with ITDs inserted furthest downstream into exon 15) (**Fig. 2B**). Since specificity increases slightly with decreasing AF (data not shown), the observed decrease in sensitivity at 2% AF would not be expected to have a significant effect on overall accuracy.

Recent data indicate that the AF of *FLT3*-ITD relative to wild-type *FLT3* is more significantly associated with poor outcomes in AML compared to ITD length or insertion position.[10] Therefore, another important aspect of ITD detection is determination of its AF relative to wild-type *FLT3*. Since PCR efficiency is affected by amplicon size, opposing primer-based methods amplify large ITD-containing fragments with lower efficiency than wild-type sequences. This results in misrepresentation of the ITD in the sequencing library and inaccurate determination of *FLT3*-ITD AF in comparison to wild-type *FLT3*. However, MBC adapters utilized in AMP are ligated to all fragments prior to amplification and contain unique molecular barcodes that allow for post-sequencing unique molecular identification. This mitigates the effects of PCR bias and allows for more accurate quantification of AFs of short and medium length ITDs. For example, a cell line containing a 30bp *FLT3*-ITD was diluted into wild-type cells to artificially create 10%, 5% and 1% AFs. After performing the VariantPlex Myeloid NGS assay, we determined the AFs to be 7.3%, 3.2% and 1.0%, respectively (data not shown), demonstrating accurate AF quantification of this *FLT3*-ITD down to 1% AF.

Next, we tested the ability of the VariantPlex Myeloid assays in conjunction with the optimized analysis algorithm to detect *FLT3*-ITDs from blood and bone marrow samples (**Fig. 3**). Twenty-seven total samples were tested, including 16 known *FLT3*-ITD- positive samples, 11 known *FLT3*-ITD-negative samples, including 2 Genome In A Bottle (GIAB) negative controls. We detected 16/16 ITDs with no false positives reported, resulting in 100% accuracy of ITD detection from clinical-type blood and bone marrow samples. All of the ITDs we detected had unique insertion positions within exon 14, ranging from chr13:28608215 to chr13:28608313, and we detected ITDs with lengths ranging from 21bp to 180bp. In addition to accurately detecting *FLT3*- ITDs, the VariantPlex assays detected concomitant non-ITD variants in *FLT3* and *NPM1* in some of the samples. *FLT3* kinase domain mutations were detected in 25% (4/16) of *FLT3*-ITD-positive blood samples, and *NPM1* variants were detected in 44% (7/16) of *FLT3*-ITD-positive blood samples, consistent with previous reports.[7,10] In 1 of these samples, both a *FLT3* kinase domain mutation and an *NPM1* variant were found to coexist with a *FLT3*- ITD (sample 6).

Sample	Input (ng)	Unique starts per GSP2*	FLT3 ITD	ITD length (bp)	Insertion position	Concordance	NPM1 variant	FLT3 variant
1	50	138	Yes	45	chr13:28608293	Yes	Yes	No
2	50	166	Yes	30	chr13:28608272	Yes	Yes	No
3	50	154	Yes	39	chr13:28608242	Yes	Yes	No
4	50	157	Yes	84	chr13:28608239	Yes	Yes	No
5	55	34	Yes	36	chr13:28608215	Yes	No	No
6	50	122	Yes	24	chr13:28608259	Yes	Yes	Yes
7	80	138	Yes	48	chr13:28608265	Yes	No	Yes
8	200	191	Yes	180	chr13:28608284	Yes	No	No
9	200	169	Yes	60	chr13:28608313	Yes	Yes	No
10	200	176	Yes	21	chr13:28608284	Yes	No	Yes
11	200	202	Yes	21	chr13:28608273	Yes	Yes	No
12	200	179	Yes	57	chr13:28608281	Yes	No	Yes
13	200	191	Yes	24	chr13:28608270	Yes	No	No
14	200	199	Yes	33	chr13:28608278	Yes	No	No
15	200	166	Yes	30	chr13:28608262	Yes	No	No
16	100	258	Yes	33	chr13:28608261	Yes	No	No
17	13	38	No	-	-	Yes	No	No
18	77	195	No	-	-	Yes	No	No
19	14	118	No	-	-	Yes	No	No
20	50	145	No	-	-	Yes	No	No
21	50	202	No	-	-	Yes	No	No
22	50	152	No	-	-	Yes	No	No
23	50	211	No	-	-	Yes	No	No
24	50	176	No	-	-	Yes	No	No
25	50	171	No	-	-	Yes	No	No
26	50	104	No	-	-	Yes	No	No
27	50	91	No	-	-	Yes	No	No

*Determined post-sequencing using Archer Analysis

Figure 3. *FLT3*-ITDs are accurately detected in AML-positive blood and bone marrow samples by VariantPlex Myeloid NGS assays. The VariantPlex Myeloid assays were used for NGS-based detection of *FLT3*-ITDs in AML-positive blood samples with known *FLT3*-ITD status, previously determined by multiple methods, including capillary gel electrophoresis (CGE). Samples 26 and 27 are Genome In A Bottle negative controls. The DNA input amounts are shown and the number of start sites per GSP2 indicates library complexity. The Archer Analysis variant pipeline was used to determine *FLT3*-ITD status, length and insertion position, with CGE concordance indicated. Concomitant *FLT3* and NPM1 variants detected simultaneously with *FLT3*-ITDs are also shown.

Conclusions

VariantPlex Myeloid targeted NGS assays, in conjunction with the data-optimized analysis algorithm in Archer Analysis, enable accurate detection of *FLT3*-ITDs from genomic DNA samples extracted from clinical-type blood and bone marrow samples. These assays are based on AMP chemistry, which utilize MBC adapters for open- ended amplification from anchored gene-specific primers for complete coverage and de novo assembly of ITD-containing regions. MBC adapters ligated prior to amplification further enable unique molecule identification for accurate AF determination. Analysis of more than 2,000 in silico datasets show that Archer Analysis accurately detects *FLT3*-ITDs across the known range of ITD lengths and insertion positions down to 1% AF. Furthermore, the VariantPlex assays together with Archer Analysis detected *FLT3*-ITDs in blood and bone marrow samples with 100% accuracy. These assays simultaneously detected *FLT3* and *NPM1* variants in *FLT3*-ITD-positive samples, highlighting the ability of AMP-based NGS to detect multiple, complex mutation types in a single assay.

Additional resources

Learn more about VariantPlex Myeloid NGS assays at **www.invitae.com/variantplex-myeloid**

For Research Use Only. Not for use in diagnostic procedures.

References

1. M. Levis, FLT3 mutations in acute myeloid leukemia: what is the best approach in 2013? *Hematology.* **2013**, 220–226 (2013).
2. K. Al-Issa, A. Nazha, Molecular landscape in acute myeloid leukemia: where do we stand in 2016. *Cancer Biol Med.* **13**, 474–482 (2016).
3. K. H. Metzeler *et al.*, Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. *Blood.* **128**, 686–698 (2016).
4. J. E. Lancet, New agents: great expectations not realized. *Best Pract Res Clin Haematol.* **26**, 269–274 (2013).
5. H. Konig, M. Levis, Targeting FLT3 to treat leukemia. *Expert Opin. Ther. Targets.* **19**, 37– 54 (2015).
6. M. Nakao *et al.*, Internal tandem duplication of the flt3 gene found in acute myeloid leukemia. *Leukemia.* **10**, 1911–1918 (1996).
7. S. Kayser *et al.*, Insertion of FLT3 internal tandem duplication in the tyrosine kinase domain-1 is associated with resistance to chemotherapy and inferior outcome. *Blood.* **114**, 2386–2392 (2009).
8. Y. Alvarado *et al.*, Treatment with FLT3 inhibitor in patients with FLT3-mutated acute myeloid leukemia is associated with development of secondary FLT3-tyrosine kinase domain mutations. *Cancer.* **120**, 2142–2149 (2014).
9. S. Schnittger *et al.*, Diversity of the juxtamembrane and TKD1 mutations (exons 13-15) in the FLT3 gene with regards to mutant load, sequence, length, localization, and correlation with biological data. *Genes Chromosomes Cancer.* **51**, 910–924 (2012).
10. R. E. Gale *et al.*, The impact of FLT3 internal tandem duplication mutant level, number, size, and interaction with NPM1 mutations in a large cohort of young adult patients with acute myeloid leukemia. *Blood.* **111**, 2776–2784 (2008).