



Published in final edited form as:

*Clin Cancer Res.* 2019 January 15; 25(2): 573–583. doi:10.1158/1078-0432.CCR-18-0655.

## Classes of ITD predict outcomes in AML patients treated with FLT3 inhibitors

Gregory W. Schwartz<sup>1</sup>, Bryan Manning<sup>4</sup>, Yeqiao Zhou<sup>1</sup>, Priya Velu<sup>1</sup>, Ashkan Bigdeli<sup>6</sup>, Rachel Astles<sup>4</sup>, Anne W. Lehman<sup>4</sup>, Jennifer J.D. Morrisette<sup>1,6</sup>, Alexander E. Perl<sup>4</sup>, Mingyao Li<sup>5</sup>, Martin Carroll<sup>4</sup>, and Robert B. Faryabi<sup>1,2,3,6</sup>

<sup>1</sup>Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

<sup>2</sup>Department of Cancer Biology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

<sup>3</sup>Abramson Family Cancer Research Institute, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

<sup>4</sup>Division of Hematology and Oncology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

<sup>5</sup>Department of Biostatistics and Epidemiology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

<sup>6</sup>Center for Personalized Diagnostics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

### Abstract

**Purpose:** Recurrent internal tandem duplication (ITD) mutations are observed in various cancers including acute myeloid leukemia (AML), where ITD mutations in tyrosine kinase receptor FLT3 are associated with poor prognostic outcomes. Several FLT3 inhibitors (FLT3i) are in clinical trials for high-risk *FLT3*-ITD-positive AML. However, the variability of survival following FLT3i treatment suggests that the mere presence of *FLT3*-ITD mutations might not guarantee effective clinical response. Motivated by the heterogeneity of *FLT3*-ITD mutations, we investigated the effects of *FLT3*-ITD structural features on the response of AML patients to treatment.

**Experimental Design:** We developed the HeatITup (HEAT diffusion for Internal Tandem dUPLICATION) algorithm to identify and quantitate ITD structural features including nucleotide

---

**Corresponding author:** Robert B. Faryabi, faryabi@penmedicine.upenn.edu, Address: Room 553 BRB II/III, 421 Curie Boulevard, Philadelphia, PA 19104-6160, Phone number: (215)573-8220.

Authors Contributions:

R.B.F., M.C., and G.W.S. designed experiments. G.W.S. and R.B.F. designed the algorithm. G.W.S. implemented the software, and performed experiments. G.W.S., A.B. obtained and organized the sequencing data. R.B.F., G.W.S., Y.Z., P.V., B.M., R.A., A.W.L., J.J.D.M., A.E.P., and M.C. collected, organized and curated electronic health record data. R.B.F., M.L. and M.C. provided expert advice. G.W.S. and R.B.F. wrote the manuscript with comments from all authors. R.B.F. supervised the project.

**Conflict-of-Interests Disclosures:** A.E. Perl is a consultant/advisory board member for Arog, Astellas, Daiichi Sankyo, Novartis, and Pfizer. No potential conflicts of interest were disclosed by the other authors.

composition. Using HeatITup, we studied the impact of ITD structural features on the clinical response to FLT3i and induction chemotherapy in *FLT3*-ITD-positive AML patients.

**Results:** HeatITup accurately identifies and classifies ITDs into newly defined categories of “typical” or “atypical” based on their nucleotide composition. A typical ITD’s insert sequence completely matches the wildtype *FLT3* whereas an atypical ITD’s insert contains nucleotides exogenous to the wildtype *FLT3*. Our analysis shows marked divergence between typical and atypical ITD mutation features. Furthermore, our data suggest that AML patients carrying typical *FLT3*-ITDs benefited significantly more from both FLT3i and induction chemotherapy treatments than patients with atypical *FLT3*-ITDs.

**Conclusions:** These results underscore the importance of structural discernment of complex somatic mutations such as ITDs in progressing towards personalized treatment of AML patients, and enable researchers and clinicians to unravel ITD complexity using the provided software.

## Introduction:

Several recurrent somatic genetic alterations in acute myeloid leukemia (AML) have been incorporated in the standard-of-care recommendations for diagnosis, response assessment, and treatment outcome prediction<sup>1, 2</sup>. Internal tandem duplication (ITD) mutations in the Fms-like tyrosine kinase 3 (*FLT3*) gene are among the most frequent somatic alterations in AML<sup>3</sup>. *FLT3*-ITDs are always in-frame insertions and result in the loss of auto-inhibitory function, leading to increased tyrosine kinase activity. The presence of *FLT3*-ITD, regardless of the cytogenetic classification, is associated with significantly increased relapse risk and decreased overall survival<sup>4-12</sup>. Hence, screening for the presence of the *FLT3*-ITD mutation is recommended<sup>4, 8, 9</sup> and genetic tests to identify *FLT3*-ITD mutations are now part of routine diagnostic workups for both *de novo* and relapsed AML patients. Recently, next-generation sequencing (NGS) diagnostic assays have been developed to test for the presence of mutant *FLT3*-ITD in AML patients<sup>13, 14</sup>. Although these new technologies enable a more detailed analysis of complex *FLT3*-ITD mutations, we still lack an understanding of the heterogeneity of ITD structures and its impact on the clinical outcomes. We present the first study to collate the effect of structural properties of *FLT3*-ITDs beyond the length of the duplicated segment itself on response to targeted therapy.

The prevalence of prognostically high-risk *FLT3*-ITD mutations resulted in the introduction of FLT3 inhibitory agents as therapeutic options for *FLT3*-ITD-positive AML. More recently, FLT3 tyrosine kinase inhibitory small-molecules (FLT3i) with various levels of specificity and potency have been developed<sup>15, 16</sup>. FLT3 inhibitors are classified based on their mechanism of interaction with the receptor<sup>15</sup> into type I (e.g. Gilteritinib (ASP2215)<sup>17, 18</sup> and Midostaurin<sup>19, 20</sup>) or type II (e.g. Sorafenib<sup>21-23</sup> Quizartinib (AC220)<sup>24-27</sup>). Although the newer generation of FLT3 inhibitors provides stronger and more prolonged anti-leukemic activity, the complete response rate still is at most 50% (depending on response criteria)<sup>17-20, 27-31</sup>, demonstrating that much remains to be learned before effective use of these agents for targeted therapy in *FLT3*-ITD-positive patients. The heterogeneity of ITD mutations could be one of the factors leading to variability in treatment response of the *FLT3*-ITD-positive AML patients.

Unlike many other hotspot mutations, ITD mutations are diverse and complex genetic alterations<sup>32, 33</sup>. While in-frame ITDs are mainly in exon 14 of *FLT3*, they exhibit marked structural heterogeneity. For example, diversity in the ITDs duplicated sequence length, duplication position, or mutation burden is reported<sup>34</sup>. Earlier studies using fluorescence-based capillary gel electrophoresis (CGE) suggest that the ITD length could impact the prognostic outcome of AML patients treated with chemotherapy-based regimens<sup>12, 32</sup>. However, the link between the response to FLT3i and *FLT3*-ITD structural features remains unknown. We hypothesized that the *FLT3*-ITD structural features beyond the length impacts clinical response to FLT3i treatment. A prerequisite to test this hypothesis was an analytical method to characterize the complexity of ITD somatic mutations from NGS readouts.

To address the unmet need for accurate identification of ITD mutations from the NGS readouts and further characterize their complexity, we developed HeatITup (HEAT diffusion for Internal Tandem dUPlication). HeatITup is an efficient and easy-to-use open-source software that is available for download from <https://github.com/faryabib/HeatITup>. Based on the divergence in the nucleotide composition of ITD mutations beyond the duplicated segment, we introduced two distinct classes of “typical” and “atypical” ITD mutations. HeatITup classifies an ITD mutation as typical if the insertion sequence completely matches the wildtype *FLT3* sequence, otherwise the ITD was classified as atypical. We applied HeatITup to samples from both the Cancer Genome Atlas (TCGA) cohort and AML patients sequenced at the University of Pennsylvania (UPENN) to identify and annotate *FLT3*-ITD mutations.

Equipped with our refined ITD classification, we systematically studied the impact of the presence of *FLT3*-ITD mutations and their structural features on the clinical response to FLT3 inhibitory agents and induction chemotherapy. Our data suggest that AML patients carrying typical ITD mutations with insert sequences entirely from the *FLT3* wildtype sequence benefit significantly more from FLT3i than patients with atypical ITDs whose inserts consist of nucleotides exogenous to the wildtype *FLT3*. Furthermore, we observed on longer overall survival (OS) in individuals harboring long typical versus long atypical *FLT3*-ITD mutations, when patients only received induction chemotherapy. Together, our results suggest that the selectivity of response to treatments depends not only on the presence of an ITD mutation in the AML genome but could also be a function of the ITD structure beyond the mutation length.

## Materials and Methods:

### The Cancer Genome Atlas (TCGA) AML cohort:

Sequencing files for 321 samples from 149 AML patients in the TCGA cohort were downloaded from the Genome Data Commons (<https://gdc.cancer.gov>)<sup>3</sup>.

### DNA sequencing:

The protocol for this study was approved by the University of Pennsylvania Medical School Institutional Review Board. *FLT3* sequencing was performed via targeted next-generation sequencing (HUP-HemeV2; the University of Pennsylvania, Philadelphia, PA) on the MiSeq

sequencing system (Illumina) as described before<sup>13</sup>. HUP-HemeV2 is a Clinical Laboratory Improvement Amendment (CLIA)-confirmed NGS assay designed at the University of Pennsylvania to detect somatic mutations in 68 genes including *FLT3*<sup>13</sup>. DNA was quantified using a fluorescent based measurement (Qubit, Life Technologies) and 50 to 250 ng of DNA was used for enrichment. The minimum DNA concentration for assay was 10 ng/uL, for a minimum of 50 ng. Following library preparation with the TruSeq Amplicon assay (Illumina), libraries were pooled and sequenced on the MiSeq system to an average depth of coverage greater than 1,000x.

### The University of Pennsylvania (UPENN) cohort:

537 amplicon-based NGS clinical samples from 431 AML patients were collected and analyzed to identify and characterize *FLT3*-ITD mutations. These samples were obtained from 2013 to 2017. The data collection was frozen in June 2017. In this cohort, 68 patients were sequenced more than once throughout the course of their disease. Overall, 139 *FLT3*-ITD-positive samples from 97 individuals were identified.

Date of diagnosis, date of NGS testing, date of censor (death or last recorded follow-up), sex, age, allogeneic transplant, FLT3i type<sup>15</sup> (type I, type II, and both) and disease status at the time of FLT3i treatment initiation were obtained for the 97 *FLT3*-ITD-positive patients from the University of Pennsylvania electronic health record (EHR) system and verified using Leukemia Tissue bank database of the University of Pennsylvania. The FLT3 Type I inhibitory agents included Gilteritinib (32), Midostaurin (8); while the Type II inhibitors included Quizartinib (14), Sorafenib (31), PLX3397 (9). 84 out of 97 patients who had complete retrievable clinical annotations and conformed to stringent clinical annotation criteria, described below, were selected to minimize clinical variability across the individuals included in our retrospective survival analysis (Table 1).

Accordingly, 68 out of 84 *FLT3*-ITD-positive patients who received a FLT3 inhibitor as part of their treatment were used to assess the correlation between *FLT3*-ITD mutation classification and survival for FLT3i-treated patients (FLT3i-treated cohort) (Table 1). These 68 FLT3i-treated patients were annotated as either newly diagnosed or relapsed AML. The newly diagnosed group consisted of individuals diagnosed with *de novo* AML with no previous history of leukemia who received induction chemotherapy, were treated with FLT3i, and were tested with the NGS-based panel within 30 days of the diagnosis date (n = 42). Alternatively, the relapsed group consisted of patients with relapsed AML who only received induction chemotherapy but not FLT3i prior to the relapse. This group was treated with both FLT3i and induction chemotherapy, and tested for FLT3 mutation by the NGS-based panel within 30 days after the relapse date (n = 26).

Most of the *FLT3*-ITD-positive AML patients at UPENN receive FLT3i treatment. So, to investigate the relationship between the *FLT3*-ITD classification and overall survival of patients who were only treated with induction chemotherapy and not FLT3i, we combined the remaining 16 out of 84 *FLT3*-ITD-positive newly diagnosed (*de novo* AML with no previous history of leukemia) patients in UPENN cohort who only received standard chemotherapy with the 35 *FLT3*-ITD-positive *de novo* AML patients from the TCGA cohort who were also treated with standard chemotherapy (non-FLT3i-treated cohort). Similar to

the newly diagnosed FLT3i-treated cohort, only the UPENN patients who received standard chemotherapy but not FLT3i and were tested for the presence of *FLT3* mutation within 30 days of the diagnosis date were included in the non-FLT3i-treated cohort (n= 51). 56 out of the total of 58 newly diagnosed FLT3-ITD-positive patients (42 FLT3i-treated and 16 chemotherapy only) were cytogenetically characterized and classified according to Medical Research Council (MRC) classification system<sup>4,35</sup>, among them none were favorable, 49 were intermediate, 6 were unfavorable, and 1 had no growth.

#### ITD Clonal assignment:

All reads with a similar *FLT3*-ITD mutation were assigned to a clone with an allele frequency (AF). We provided a helper program to assist in grouping together similar mutations into clones. We only considered clones with an AF  $\geq 1\%$ , resulting in 189 typical and 50 atypical clones in 97 *FLT3*-ITD-positive individuals in the UPENN cohort. Due to an  $\sim 100$ -fold lower sequencing coverage of the TCGA samples, we only included clones if their *FLT3*-ITD duplication length was 15 or greater. Each individual's representative *FLT3*-ITD clone was determined based on the dominant clone from the earliest sample in the UPENN cohort or the longest duplication in the TCGA cohort due to the significantly less coverage of these samples. See Supplementary Information for further description.

#### Statistical Analysis:

The Mann-Whitney U test was used in comparing duplicated and spacer segment lengths, as well as purine and pyrimidine abundance. Exogenous nucleotide abundance was compared using Tukey's HSD after ANOVA. The Chi-squared test was used to assess the point mutation co-occurrence frequency. Overall survival (OS), defined as the time from diagnosis until death or last follow-up, was determined using Kaplan-Meier survival curves. Survival statistical analyses were performed using Cox regressions<sup>36</sup>. Univariate Cox regressions were used to select possible confounding variables among the following features: disease stage (newly diagnosed and relapsed), sex, age, allogenic transplant, FLT3i type, disease status at the time of FLT3i treatment initiation, and clone frequency. Wherever applicable, the significant confounding variables were controlled for in the subsequent multivariate Cox regression models of ITD classification. The proportional hazards assumption for each variable was tested and met using Schoenfeld residuals<sup>37</sup>. Adjusted p-values for the variables of interest (FLT3i treatment, duplication length, ITD classification, and spacer sequence exogenous ratio) were determined by Benjamini-Hochberg procedure. Wherever applicable, the entire spectrum of the ITD spacer sequence exogenous ratio values was used – the exogenous ratio cutoff in figures was used for visualization purposes only.

#### Results:

Unlike many other hotspot mutations, ITDs are diverse and complex genetic alterations. We observed that the inserted segment of DNA could appear immediately next to its origin of duplication or there could be intervening nucleotides (Figures 1a and 3). While ITD mutations are mainly in the *FLT3* exon 14, the insertion location and sequences, the length of the duplicate segment, and the composition of the DNA segment between the duplication and its origin add to the ITD structural complexity (Figures 1a and 3).

We defined several attributes to capture ITD diversity. We termed the part of an inserted DNA segment that is a duplication of the reference as “duplication”, and the nucleotides in-between the duplication and its origin as “spacer” (Figure 1a). We observed that a spacer could consist of nucleotides that were either endogenous to the *FLT3* wildtype sequence or were not part of the wildtype exon (Figures 1a and 3). To this end, we classified an ITD as “typical” if the insertion sequence completely matched the *FLT3* wildtype sequence, otherwise the ITD was classified as “atypical” (Figures 1a and 3). As such, the difference between typical and atypical ITDs is based on the presence of nucleotides, excluding point mutations, that are exogenous to *FLT3*-wildtype sequences in an ITD insertion. Importantly, while atypical and typical ITDs could have spacer sequences, they both resulted in in-frame mutations.

Since ITD mutations are a combination of alterations, commonly used tools such as Pindel<sup>38</sup> do not directly detect them. We developed HeatITup to address the unmet need for accurate identification of ITD mutations from NGS short sequence reads, and further characterize the complexity of ITD mutations based on the aforementioned structural attributes. Rather than search for an insertion as in<sup>3, 38–40</sup>, HeatITup can analyze both mapped (heatitup program) and unmapped (heatitup-complete program) short-read sequences and detects the largest duplicated sequences; characterizes the sequence composition of the insertion; and classifies an ITD as typical or atypical (Figures 1a and 3, see Supplemental Information). By repeated usage of suffix trees with “end mutations” (see Materials and Methods, and Supplemental Information), HeatITup is able to detect both the duplications as well as any possible point mutations within them (Figures 1b and 1c). HeatITup also analyzes the composition of spacer sequences by using a combination of Hamming distance comparisons with a reference sequence and discrete Gaussian kernel-based heat diffusion to detect the sequence segment within the spacer that may be exogenous to the gene wildtype sequence (Figures 1b and 1c, Materials and Methods, and Supplemental Information). As such, HeatITup is the first method capable of annotating duplication, spacer, and the exogenous segments of the insertion in addition to detecting ITDs.

To assess the accuracy of HeatITup, we first performed a comparative analysis against four commonly used methods based on 321 samples from 149 individuals profiled in the TCGA AML cohort using capture-based whole exome sequencing<sup>3</sup>. These samples were previously scored for the presence of ITD mutations in *FLT3* by the method used in the original analysis of the TCGA AML cohort<sup>3</sup>, ITD Assembler<sup>41</sup>, Pindel<sup>38</sup>, and Genomen<sup>42</sup>. HeatITup found 7 new *FLT3*-ITD-positive individuals that were manually verifiable but missed by these methods (Figure 2). Among the 35 detected *FLT3*-ITD-positive individuals, 20 harbored typical and the remaining carried atypical ITDs. HeatITup’s heatitup program (alignment-based, see Supplementary Information) but not heatitup-complete program (alignment-free, see Supplementary Information) missed 1 *FLT3*-ITD-positive individual who was only identified by the ITD Assembler algorithm. Compared to the earlier analysis, HeatITup increased the percentage of detected *FLT3*-ITD-positive cases in the TCGA cohort from 19% to 23%, which is closer to the expected frequency of *FLT3*-ITD-positive cases in a representative AML cohort based on other methodologies such as CGE, while did not falsely identified any sample as *FLT3*-ITD mutated. Together, these data exhibit that HeatITup accurately detects ITD mutations while providing high sensitivity.

We then analyzed amplicon-based NGS of 216 patients in the UPENN cohort to assess the accuracy of *FLT3*-ITD mutation detection. We compared our results with a previous analysis of these samples at the UPENN clinical sequencing laboratory where a two-step CLIA-confirmed analysis workflow, consisting of a GATK/Pindel-based pipeline followed by manual variant review, were used (see Materials and Methods). HeatITup identified 113 *FLT3*-ITD-positive samples from 71 individuals (Figure 2b). Our analysis improved the detection of *FLT3*-ITD mutations and found an additional 5 *FLT3*-ITD-positive individuals that were manually verifiable but reported as *FLT3*-WT by the UPENN CLIA-confirmed workflow.

We next assessed the accuracy of HeatITup to classify samples as typical or atypical. To this end, we independently analyzed 32 UPENN samples with HeatITup and experienced genetic reviewers who verified the presence and subsequently classified ITD mutations. When the HeatITup and manual detection/classification results were compared, 38 out of 39 ITDs were classified concordantly while one sample, CPDC140113, had a discordant classification (Figure 3, c.f. red-boxed individual). Upon further inspection, we found that the misclassified *FLT3*-ITD in sample CPDC140113 was indeed atypical and HeatITup was able to detect and correct the error of manual classification (Table S1). Together, these evaluations exhibited the precision of HeatITup in detecting and classifying *FLT3*-ITD mutations into typical or atypical categories while providing high sensitivity.

We quantified the clonality and structural characteristics of 139 *FLT3*-ITD-positive samples from 97 individuals in the UPENN cohort to assess the structural heterogeneity of ITD mutations (see Materials and Methods). In our cohort, typical and atypical *FLT3*-ITD clones occurred with 79.7% and 20.3% frequencies, respectively. We found that our cohort consisted of similar proportions of mono- and polyclonal cases (Figure 4a). While 46.4% of individuals harbored one *FLT3*-ITD clone, 7.2% of patients contained more than 3 clones (Figure 4a). Surprisingly, 45.4% of polyclonal individuals were homogeneous for the ITD classification resulting in 91.8% of patients harboring a single class of *FLT3*-ITD (Figure 4b), which may point to a single ongoing mechanism of mutation even in the polyclonal samples. These observations suggest potentially distinct mechanisms in the etiology of the two classes of *FLT3*-ITD mutations or that they are enriched in particular genetic backgrounds.

As typical and atypical clones may have other distinct structural features beyond the exogenous nucleotides within the insertion, we sought to quantify these characteristics. We observed marginally significant longer duplications (Individuals:  $p = 0.075$ , Figure 4c; All clones:  $p = 0.3$ , Figure S1a) and significantly longer spacers (Individuals:  $p < 2.22e-16$ , Figure 4d; All clones:  $p < 2.22e-16$ , Figure S1b) in atypical versus typical clones, which may point to the heterogeneity of these longer duplications in earlier studies<sup>33, 43, 44</sup>.

We further characterized the exogenous region of ITD mutations. While the median length of the exogenous segment was 3 bp, we observed a clone with 51 exogenous nucleotides in our cohort (Figures 4e and S1c). Upon further investigation, we found this segment originated from the intron 3' of the *FLT3* exon 14. While we were able to clarify the origin of the exogenous segment of the ITD in this case, for the most part, smaller exogenous

segments in other ITDs were highly promiscuous across the genome and it was challenging to resolve their origin. As such, we only characterized the nucleotide makeup of inserted segments to gain insight into biases of their nucleotide usage, and observed that the exogenous segments were significantly enriched for purines ( $p = 0.01$ ) and even more significantly for guanine nucleotides ( $G > A$  ( $p = 0.06$ ),  $G > C$  ( $p < 1e-4$ ),  $G > T$  ( $p = 0.005$ ), and ( $p > 0.05$ ) for all others, Figure S1d).

As the prognostic effect of individual mutations could be impacted by the presence of other mutations, we investigated co-occurring point mutations in *FLT3*-ITD-positive genomes. Among 68 genes interrogated by the HUP-HemeV2 panel, there was no significant difference between the mutation patterns of typical and atypical *FLT3*-ITD cases (Chi-squared test:  $p > 0.05$ ). The distribution of known AML driver mutations (*DNMT3A*, *ASXL1*, *IDH1*, *IDH2*, and *TET2*<sup>45</sup>) was also not significantly different between the *FLT3*-ITD classes. Earlier studies suggest that mutation in *FLT3* tyrosine kinase domain (TKD) may impact the FLT3i response<sup>16, 18, 46</sup>. The co-occurrence of *FLT3*-ITD and TKD mutations was rare in our *FLT3*-ITD-positive cohort (10.3%), as expected<sup>20, 47, 48</sup>, and was not enriched in any of the *FLT3*-ITD classes. Together, these analysis suggest that co-occurring point mutations are not differentially associated with typical or atypical *FLT3*-ITD cases.

Observing a lack of any significant co-occurring mutation and a marked divergence of typical and atypical ITDs (Figures 4 and S1), we investigated whether *FLT3*-ITD structural features impact survival rate of AML patients. To this end, we used univariate Cox regressions to identify possible confounding variables among disease stage (newly diagnosed and relapsed), sex, age, allogenic transplant, disease status at the time of FLT3i treatment initiation, clone frequency, and FLT3i type; and identified the allogenic transplant (Cox PH: HR = 0.440, CI = 0.256 – 0.757,  $p = 0.00301$ ) and the disease status at the time of FLT3i treatment initiation (Cox PH: HR = 4.35, CI = 1.54 – 12.2,  $p = 0.00539$ ) as candidate variables to consider along with the ITD classification for further multivariate survival analysis (Table S2). There have been conflicting reports on whether *FLT3*-ITD duplication length is a significant prognostic factor in AML patients treated with standard chemotherapy. In some reports longer duplication length is associated with inferior OS<sup>33, 44</sup>, while other studies did not find any link between the ITD length and OS of AML patients after standard chemotherapy<sup>32, 49</sup>. The correlation between ITD duplication length and OS was insignificant in our *FLT3*-ITD-positive patients treated with standard chemotherapy (Table 1, see the non-FLT3i-treated cohort in Materials and Methods, Table S3). The lack of correlation between duplication length and survival incentivized the search for alternative sequence characteristics that could describe the heterogeneity of survival. As such, we inquired whether the appearance of exogenous nucleotides in ITD mutations impacted OS of patients in the non-FLT3i-treated cohort. Considering ITD classification alone, we observed on average longer OS in patients harboring typical versus atypical *FLT3*-ITD mutations, however this difference was not statistically significant (Tables S4 and S5). Given the heterogeneity in the ITD lengths, we next looked at the top third longest duplication lengths in our non-FLT3i-treated cohort and found that individuals with typical *FLT3*-ITDs had 77.4% reduced hazard compared to the patients with atypical ITDs (Cox PH: HR = 0.226, CI = 0.0764 – 0.669,  $p = 0.00722$ , Figure 5a, Table S6). Observing that atypical ITD



adversely impacted OS, we examined the relationship between the fraction of exogenous nucleotides in *FLT3*-ITD spacers and OS in the non-FLT3i-treated cohort. Higher fraction of exogenous nucleotides alone had no significant correlation with OS in the non-FLT3i-treated cohort (Table S7), but exhibited a 159 times increased hazard when the top third longest duplication lengths were considered (Cox PH: HR = 159.0, CI = 4.92 – 5160,  $p = 0.00427$ , Figure 5b, Table S8). Together, these data suggest that atypical ITD mutations may lower the overall survival of AML patients treated with standard chemotherapy.

We next tested whether the ITD classification influences the overall survival of 68 FLT3i-treated patients with complete clinical annotations (Table 1, see the FLT3i-treated cohort in Materials and Methods). In our cohort, FLT3i treatment significantly improved the overall survival of the *FLT3*-ITD-positive patients, as expected<sup>16, 20, 25, 26, 50</sup> (Cox PH: HR = 0.0432, CI = 0.00583 – 0.320,  $p = 0.00209$ , Table S9). Upon verifying the benefit of FLT3i treatment, we tested the impact of *FLT3*-ITD structural features on OS of the FLT3i-treated patients. We observed that duplication length had no significant effect on the FLT3i-treated AML patient OS (Table S10). When segregating the cohort into typical and atypical *FLT3*-ITDs, we found that carrying a typical ITD genotype reduced hazard by 66.3% (Cox PH: HR = 0.337, CI = 0.165 – 0.689,  $p = 0.00286$ , Figure 5c, Table S11), suggesting in the FLT3i-treated cohort typical patients had superior OS. When tested for the impact of exogenous nucleotide presence in the ITDs, we observed that an increased fraction of the *FLT3*-ITD spacer exogenous sequences was positively correlated with a lower OS in the FLT3i-treated cohort, for all and not just long duplication lengths (Cox PH: HR = 4.96, CI = 1.33 – 18.5,  $p = 0.0173$ , Figure 5d, Table S12). Taken together, these data suggest that the correlation of *FLT3*-ITD structure with prolonged survival in response to treatments could be a multivariate relationship, potentially dependent on the length and the composition of inserted sequences. Nevertheless, our observations suggest that the selectivity of response to FLT3i depends not only on the presence but also the nucleotide composition of ITD mutations, and AML patients with atypical *FLT3* mutations on average may have inferior OS in response to treatment with standard chemotherapy as well as FLT3 inhibitory agents.

## Discussion:

In this study, we introduce two distinct classes of typical and atypical ITDs based on the mutation nucleotide composition, show significant divergence between the characteristics of these two ITD classes, and develop an efficient and easy-to-use software (available at <https://github.com/faryabib/HeatITup>) to classify and annotate ITD mutations. Notably, our data show that responses to FLT3i treatment correlates with the classification and structural features of *FLT3*-ITD mutations, and suggest that carrying atypical ITD mutations with nucleotides mainly exogenous to the *FLT3* wildtype sequence could lower the survival rate in AML patients. We thus propose that the complexity of ITD mutations should be characterized and considered to inform clinical decision-making and therapeutic algorithms for AML patients. To this end, we have packaged HeatITup as an open-source tool with the aim of enabling similar analysis in larger independent cohorts that could further elucidate the potential effects of heterogeneity in ITD mutations. Although this study is a step towards implementing precision medicine for AML patients, more comprehensive studies are warranted to fully understand the complexity of response to FLT3i treatments. Our data

underscore the importance of understanding *FLT3*-ITD mutation complexity and interactions between *FLT3* and other recurrent mutations in AML patients to improve the prediction of response to FLT3i-based therapies.

All the typical and atypical *FLT3*-ITD mutations identified in our study disrupt the *FLT3* exon 14 sequence coding for the FLT3 juxtamembrane domain (JM). While the crystal structure of FLT3 suggests that ITD mutations negate the intrinsic auto-inhibitory activity of the JM domain <sup>9, 43</sup>, it is unclear how highly heterogeneous characteristics of ITDs affect receptor function. One could speculate that more complex ITD mutations deviate further from the receptor wildtype structure. Long atypical ITDs with large fraction of exogenous nucleotides might further disrupt the auto-inhibitory activity of the JM domain, while short duplications or less deviation from the endogenous sequences may preserve some of the intrinsic auto-inhibitory property of the receptor. Although our retrospective study suggests that the *FLT3*-ITD structure could play a role in determining patient outcomes, we do not provide data about the changes in protein structure due to typical or atypical ITD mutations; hence, further studies, such as <sup>15, 46, 51</sup>, are essential to clarify the mechanisms of the observed effects and elucidate whether the diversity of ITD mutations differentially impacts FLT3 receptor activity and drug response.

Recurrent ITD mutations are not unique to *FLT3* and have been reported in a related receptor KIT <sup>9</sup> and histone-lysine N-methyltransferase 2A (also known as mixed lineage leukemia) (KMT2A/MLL1) in acute lymphocytic leukemia <sup>52</sup>. Also, ITD mutations are frequently observed in solid tumors such as BCOR in clear cell sarcoma of the kidney <sup>53</sup>. Our initial results warrant investigation of the link between the complexity of recurrent ITD mutations in genes beyond *FLT3* in AML and other cancers. Our algorithm, HeatITup, provides an accessible computational solution for accurate detection and classification of ITD that will enable these studies.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgment:

We are grateful to Drs. Michael D. Feldman, Kojo Elenitoba-Johnson and Warren Pear for their advice. We thank the staff at the Center for Personalized Diagnostics at the University of Pennsylvania for their support. This work was supported in part by funds from the Penn Precision Acceleration Fund (to R.B. Faryabi) and T32-CA009140 (to G.W. Schwartz).

## Abbreviations:

<b>AML</b>	acute myeloid leukemia
<b>ITD</b>	internal tandem duplication
<b>HeatITup</b>	HEAT diffusion for Internal Tandem dUPLICATION
<b>NGS</b>	next-generation sequencing
<b>FLT3i</b>	FLT3 tyrosine kinase inhibitory small-molecule

<b>CLIA</b>	Clinical Laboratory Improvement Amendment
<b>EHR</b>	electronic health record
<b>OS</b>	Overall survival

## References:

1. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127: 2391–2405. [PubMed: 27069254]
2. Vardiman JW, Harris NL, Brunning RD. The World Health Organization (WHO) classification of the myeloid neoplasms. *Blood*. 2002;100: 2292–2302. [PubMed: 12239137]
3. Cancer Genome Atlas Research N, Ley TJ, Miller C, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*. 2013;368: 2059–2074. [PubMed: 23634996]
4. Dohner H, Estey EH, Amadori S, et al. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood*. 2010;115: 453–474. [PubMed: 19880497]
5. Gregory TK, Wald D, Chen Y, Vermaat JM, Xiong Y, Tse W. Molecular prognostic markers for adult acute myeloid leukemia with normal cytogenetics. *J Hematol Oncol*. 2009;2: 23. [PubMed: 19490647]
6. Kayser S, Schlenk RF, Londono MC, et al. Insertion of FLT3 internal tandem duplication in the tyrosine kinase domain-1 is associated with resistance to chemotherapy and inferior outcome. *Blood*. 2009;114: 2386–2392. [PubMed: 19602710]
7. Kiyoi H, Naoe T, Nakano Y, et al. Prognostic implication of FLT3 and N-RAS gene mutations in acute myeloid leukemia. *Blood*. 1999;93: 3074–3080. [PubMed: 10216104]
8. Kottaridis PD, Gale RE, Frew ME, et al. The presence of a FLT3 internal tandem duplication in patients with acute myeloid leukemia (AML) adds important prognostic information to cytogenetic risk group and response to the first cycle of chemotherapy: analysis of 854 patients from the United Kingdom Medical Research Council AML 10 and 12 trials. *Blood*. 2001;98: 1752–1759. [PubMed: 11535508]
9. Levis M, Small D. FLT3: ITDoes matter in leukemia. *Leukemia*. 2003;17: 1738–1752. [PubMed: 12970773]
10. Meshinchi S, Alonzo TA, Stirewalt DL, et al. Clinical implications of FLT3 mutations in pediatric AML. *Blood*. 2006;108: 3654–3661. [PubMed: 16912228]
11. Schlenk RF, Dohner K, Krauter J, et al. Mutations and treatment outcome in cytogenetically normal acute myeloid leukemia. *N Engl J Med*. 2008;358: 1909–1918. [PubMed: 18450602]
12. Thiede C, Studel C, Mohr B, et al. Analysis of FLT3-activating mutations in 979 patients with acute myelogenous leukemia: association with FAB subtypes and identification of subgroups with poor prognosis. *Blood*. 2002;99: 4326–4335. [PubMed: 12036858]
13. Sehgal AR, Gimotty PA, Zhao J, et al. DNMT3A Mutational Status Affects the Results of Dose-Escalated Induction Therapy in Acute Myelogenous Leukemia. *Clin Cancer Res*. 2015;21: 1614–1620. [PubMed: 25609058]
14. Spencer DH, Abel HJ, Lockwood CM, et al. Detection of FLT3 internal tandem duplication in targeted, short-read-length, next-generation sequencing data. *J Mol Diagn*. 2013;15: 81–93. [PubMed: 23159595]
15. Ke YY, Singh VK, Coumar MS, et al. Homology modeling of DFG-in FMS-like tyrosine kinase 3 (FLT3) and structure-based virtual screening for inhibitor identification. *Sci Rep*. 2015;5: 11702. [PubMed: 26118648]
16. Larrosa-Garcia M, Baer MR. FLT3 Inhibitors in Acute Myeloid Leukemia: Current Status and Future Directions. *Mol Cancer Ther*. 2017;16: 991–1001. [PubMed: 28576946]
17. Altman JK, Perl AE, Cortes JE, Levis MJ. Antileukemic Activity and Tolerability of ASP2215 80mg and Greater in FLT3 Mutation-Positive Subjects with Relapsed or Refractory Acute Myeloid

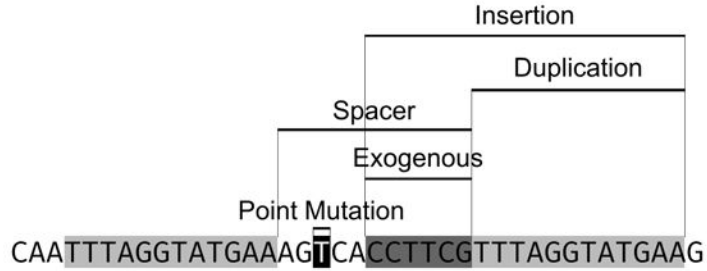
- Leukemia: Results from a Phase 1/2, Open-Label, Dose-Escalation/Dose-Response Study. ASH Annual Meeting, 2016.
18. Lee LY, Hernandez D, Rajkhowa T, et al. Preclinical studies of gilteritinib, a next-generation FLT3 inhibitor. *Blood*. 2017;129: 257–260. [PubMed: 27908881]
  19. Stone RM, DeAngelo DJ, Klimek V, et al. Patients with acute myeloid leukemia and an activating mutation in FLT3 respond to a small-molecule FLT3 tyrosine kinase inhibitor, PKC412. *Blood*. 2005;105: 54–60. [PubMed: 15345597]
  20. Stone RM, Mandrekar SJ, Sanford BL, et al. Midostaurin plus Chemotherapy for Acute Myeloid Leukemia with a FLT3 Mutation. *N Engl J Med*. 2017;377: 454–464. [PubMed: 28644114]
  21. Chen YB, Li S, Lane AA, et al. Phase I trial of maintenance sorafenib after allogeneic hematopoietic stem cell transplantation for fms-like tyrosine kinase 3 internal tandem duplication acute myeloid leukemia. *Biol Blood Marrow Transplant*. 2014;20: 2042–2048. [PubMed: 25239228]
  22. Metzelder S, Wang Y, Wollmer E, et al. Compassionate use of sorafenib in FLT3-ITD-positive acute myeloid leukemia: sustained regression before and after allogeneic stem cell transplantation. *Blood*. 2009;113: 6567–6571. [PubMed: 19389879]
  23. Metzelder SK, Schroeder T, Finck A, et al. High activity of sorafenib in FLT3-ITD-positive acute myeloid leukemia synergizes with allo-immune effects to induce sustained responses. *Leukemia*. 2012;26: 2353–2359. [PubMed: 22504140]
  24. Chao Q, Sprankle KG, Grotzfeld RM, et al. Identification of N-(5-tert-butyl-isoxazol-3-yl)-N'-{4-[7-(2-morpholin-4-yl-ethoxy)imidazo[2,1-b][1,3]benzothiazol-2-yl]phenyl}urea dihydrochloride (AC220), a uniquely potent, selective, and efficacious FMS-like tyrosine kinase-3 (FLT3) inhibitor. *J Med Chem*. 2009;52: 7808–7816. [PubMed: 19754199]
  25. Cortes JE. Final results of a phase 2 open-label, monotherapy efficacy and safety study of quizartinib (AC220) in patients  $\geq$  60 years of age with FLT3 ITD positive or negative relapsed/refractory acute myeloid leukemia. ASH Annual Meeting, 2012:Abstract 120(121):148.
  26. Levis MJ. Final results of a phase 2 open-label, monotherapy efficacy and safety study of quizartinib (AC220) in patients with FLT3-ITD positive or negative relapsed/refractory acute myeloid leukemia after second-line chemotherapy or hematopoietic stem cell transplantation.. ASH Annual Meeting, 2012:Abstract 120(121):673.
  27. Nybakken GE, Canaani J, Roy D, et al. Quizartinib elicits differential responses that correlate with karyotype and genotype of the leukemic clone. *Leukemia*. 2016;30: 1422–1425. [PubMed: 26585411]
  28. Knapper S, Burnett AK, Littlewood T, et al. A phase 2 trial of the FLT3 inhibitor lestaurtinib (CEP701) as first-line treatment for older patients with acute myeloid leukemia not considered fit for intensive chemotherapy. *Blood*. 2006;108: 3262–3270. [PubMed: 16857985]
  29. Levis M, Brown P, Smith BD, et al. Plasma inhibitory activity (PIA): a pharmacodynamic assay reveals insights into the basis for cytotoxic response to FLT3 inhibitors. *Blood*. 2006;108: 3477–3483. [PubMed: 16857987]
  30. Levis M, Ravandi F, Wang ES, et al. Results from a randomized trial of salvage chemotherapy followed by lestaurtinib for patients with FLT3 mutant AML in first relapse. *Blood*. 2011;117: 3294–3301. [PubMed: 21270442]
  31. Smith BD, Levis M, Beran M, et al. Single-agent CEP-701, a novel FLT3 inhibitor, shows biologic and clinical activity in patients with relapsed or refractory acute myeloid leukemia. *Blood*. 2004;103: 3669–3676. [PubMed: 14726387]
  32. Gale RE, Green C, Allen C, 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*. 2008;111: 2776–2784. [PubMed: 17957027]
  33. Meshinchi S, Stirewalt DL, Alonzo TA, et al. Structural and numerical variation of FLT3/ITD in pediatric AML. *Blood*. 2008;111: 4930–4933. [PubMed: 18305215]
  34. Schnittger S, Schoch C, Dugas M, et al. Analysis of FLT3 length mutations in 1003 patients with acute myeloid leukemia: correlation to cytogenetics, FAB subtype, and prognosis in the AMLCG study and usefulness as a marker for the detection of minimal residual disease. *Blood*. 2002;100: 59–66. [PubMed: 12070009]

35. Grimwade D, Hills RK, Moorman AV, et al. Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. *Blood*. 2010;116: 354–365. [PubMed: 20385793]
36. Cox DR. Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1972;34: 187–220.
37. Schoenfeld D Partial Residuals for the Proportional Hazards Regression-Model. *Biometrika*. 1982;69: 239–241.
38. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics*. 2009;25: 2865–2871. [PubMed: 19561018]
39. Fang H, Bergmann EA, Arora K, et al. Indel variant analysis of short-read sequencing data with Scalpel. *Nat Protoc*. 2016;11: 2529–2548. [PubMed: 27854363]
40. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20: 1297–1303. [PubMed: 20644199]
41. Rustagi N, Hampton OA, Li J, et al. ITD assembler: an algorithm for internal tandem duplication discovery from short-read sequencing data. *BMC Bioinformatics*. 2016;17: 188. [PubMed: 27121965]
42. Chiba K, Shiraishi Y, Nagata Y, et al. Genomon ITDetector: a tool for somatic internal tandem duplication detection from cancer genome sequencing data. *Bioinformatics*. 2015;31: 116–118. [PubMed: 25192740]
43. Griffith J, Black J, Faerman C, et al. The structural basis for autoinhibition of FLT3 by the juxtamembrane domain. *Mol Cell*. 2004;13: 169–178. [PubMed: 14759363]
44. Stirewalt DL, Kopecky KJ, Meshinchi S, et al. Size of FLT3 internal tandem duplication has prognostic significance in patients with acute myeloid leukemia. *Blood*. 2006;107: 3724–3726. [PubMed: 16368883]
45. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med*. 2016;374: 2209–2221. [PubMed: 27276561]
46. Smith CC, Wang Q, Chin CS, et al. Validation of ITD mutations in FLT3 as a therapeutic target in human acute myeloid leukaemia. *Nature*. 2012;485: 260–263. [PubMed: 22504184]
47. Smith CC, Lin K, Stecula A, Sali A, Shah NP. FLT3 D835 mutations confer differential resistance to type II FLT3 inhibitors. *Leukemia*. 2015;29: 2390–2392. [PubMed: 26108694]
48. Elyamany G, Awad M, Alsuhaibani O, et al. FLT3 Internal Tandem Duplication and D835 Mutations in Patients with Acute Lymphoblastic Leukemia and its Clinical Significance. *Mediterr J Hematol Infect Dis*. 2014;6: e2014038. [PubMed: 24959335]
49. Ponziani V, Gianfaldoni G, Mannelli F, et al. The size of duplication does not add to the prognostic significance of FLT3 internal tandem duplication in acute myeloid leukemia patients. *Leukemia*. 2006;20: 2074–2076. [PubMed: 16990788]
50. Stein EM. Molecularly targeted therapies for acute myeloid leukemia. *Hematology Am Soc Hematol Educ Program*. 2015;2015: 579–583. [PubMed: 26637775]
51. Smith CC, Lasater EA, Lin KC, et al. Crenolanib is a selective type I pan-FLT3 inhibitor. *Proc Natl Acad Sci U S A*. 2014;111: 5319–5324. [PubMed: 24623852]
52. Whitman SP, Strout MP, Marcucci G, et al. The partial nontandem duplication of the MLL (ALL1) gene is a novel rearrangement that generates three distinct fusion transcripts in B-cell acute lymphoblastic leukemia. *Cancer Res*. 2001;61: 59–63. [PubMed: 11196198]
53. Roy A, Kumar V, Zorman B, et al. Recurrent internal tandem duplications of BCOR in clear cell sarcoma of the kidney. *Nat Commun*. 2015;6: 8891. [PubMed: 26573325]

**Statement of Translational Relevance:**

ITD mutations are diverse and complex genetic alterations and result in increased kinase activity of FLT3 receptor in AML patients. Given that the presence of FLT3-ITD mutations is associated with poor prognostic outcomes, several FLT3 inhibitors (FLT3i) for targeted therapy of FLT3-ITD-positive AML patients were developed. We tested the hypothesis that the FLT3-ITD structural features beyond the length impact response to the targeted therapy. Our data suggest that the selectivity of response to FLT3i depends not only on the presence, but also the nucleotide composition of ITD mutations in newly diagnosed and relapsed AML. To enable investigating the diversity of ITD mutations, we developed an efficient and easy-to-use software to detect and characterize ITDs. Our software could inform clinical decision making for patients carrying this class of complex mutations by facilitating similar studies in AML and other malignancies with high rate of ITD mutations.

a.



Typical: CAATTTAGGTATGAATTTAGGTATGAAAGCCAGCTACAG

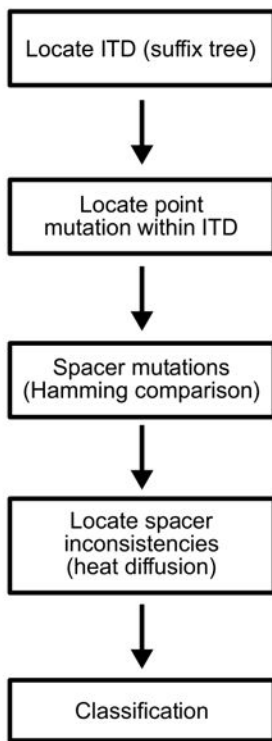
Typical: CAATTTAGGTATGAAAGCCATTTAGGTATGAAGCTACAG

Typical: CAATTTAGGTATGAAAGCCATTTAAGTATGAAGCTACAG

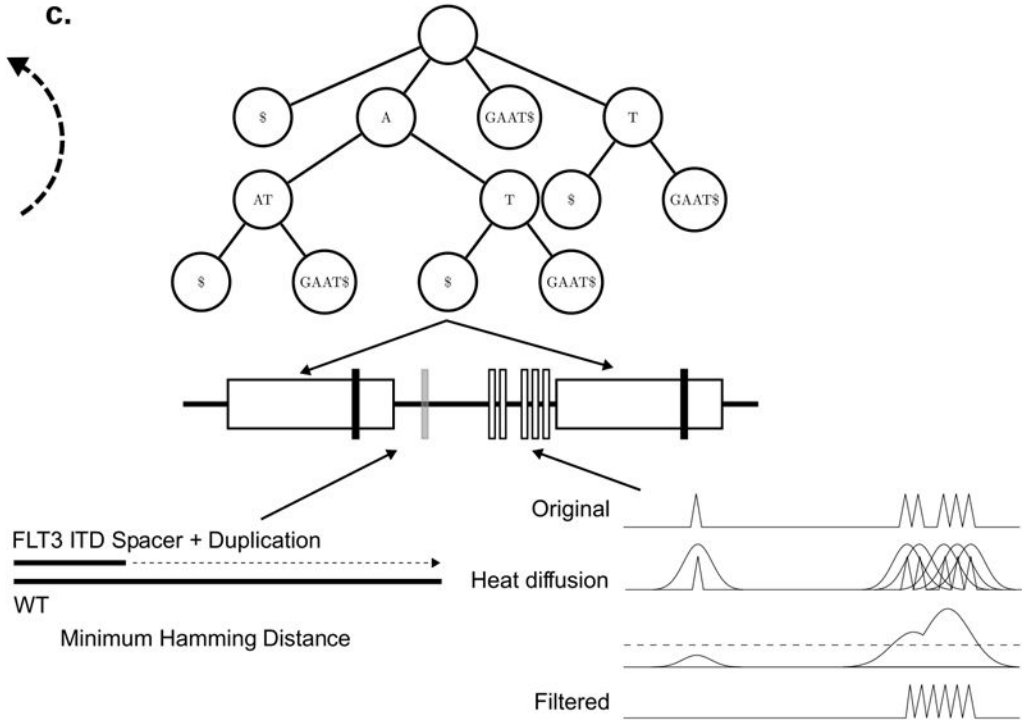
Atypical: CAATTTAGGTATGAAACCTTCGTTTAGGTATGAAAGCCAG

Atypical: CAATTTAGGTATGAAAGCCACCTTCGTTTAGGTATGAAG

b.



c.



**Figure 1. Schematic of representative typical and atypical ITD mutations and HeatITup algorithm.**

a) Schematic of representative typical and atypical ITD mutations with marked insertion, duplication, spacer, and exogenous nucleotides. First sequence: marked features of ITD structure. Remaining sequences are examples of typical and atypical ITDs. Top to bottom: Duplication with no spacer; Duplication with spacer; Duplication with spacer and a point mutation; Duplication with entire spacer exogenous to the exon sequence; Duplication with part of spacer exogenous to the exon sequence. Light gray: duplicated, black: point mutation, dark gray: exogenous sequences. b) Outline of the HeatITup algorithm. c) Schematic of HeatITup modules and their function to analyze an ITD mutation. In this

schematic, an ITD in the *FLT3* sequence (middle, black horizontal line) is analyzed using a suffix tree (top), a sliding window (bottom left), and heat diffusion (bottom right). The repeated substring (duplication, white rectangles on *FLT3* sequence) with potential point mutations (vertical black bars) is identified using repeated suffix tree with “end mutation” (top). The sliding window of minimum Hamming distance (bottom left) identifies differences within the spacer (vertical bars in the spacer). To differentiate between the exogenous nucleotides in the spacer (vertical white bars on the spacer) and the point mutations (vertical gray bar on the spacer), heat diffusion is applied to the signal processed from the spacer (bottom right).

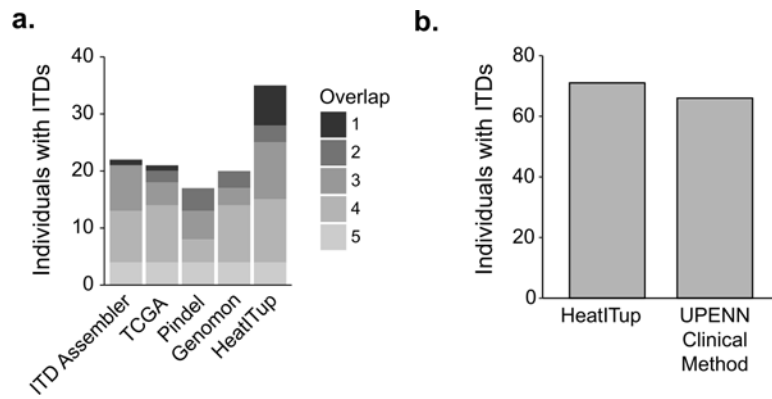
Author Manuscript

Author Manuscript

Author Manuscript

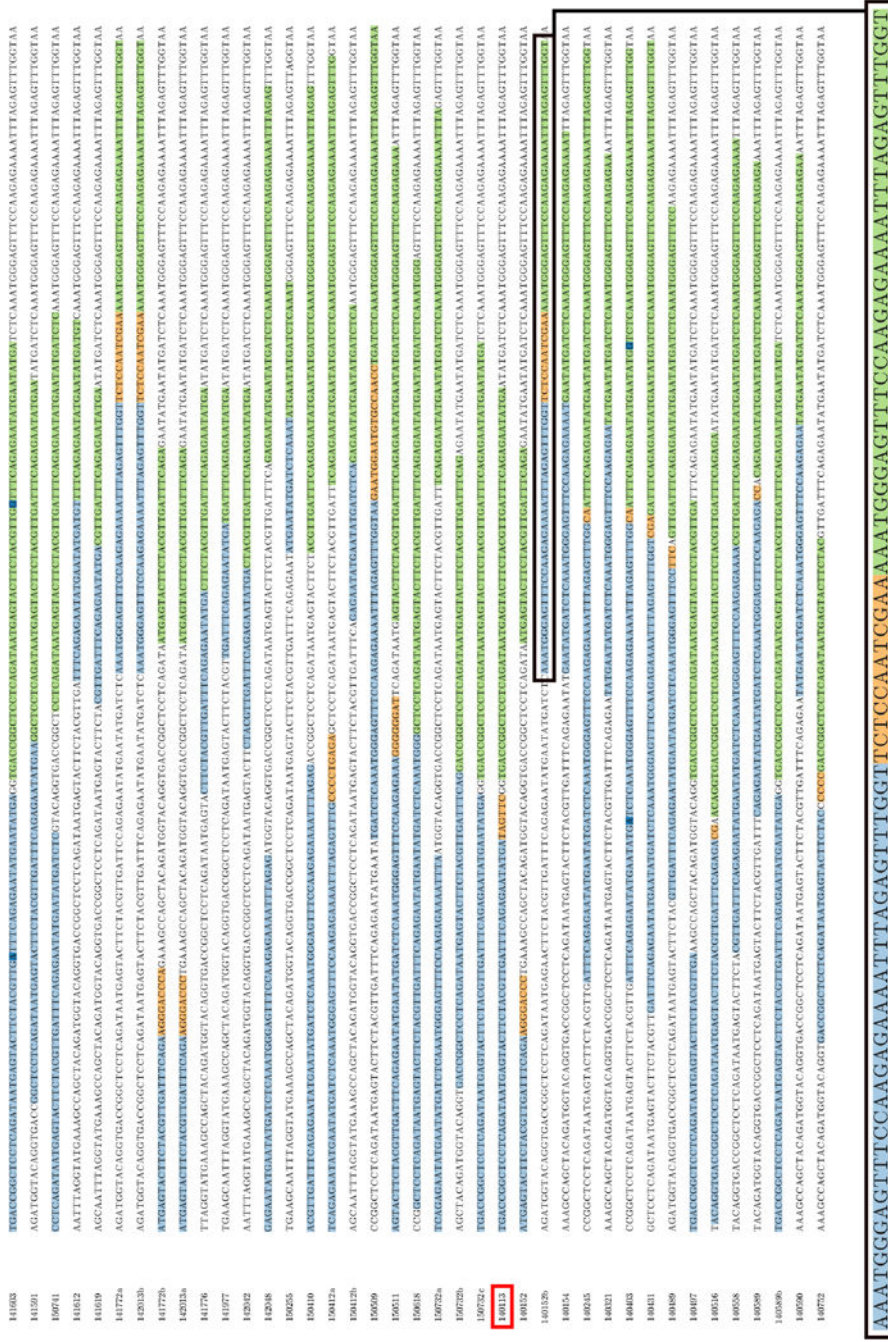
Author Manuscript





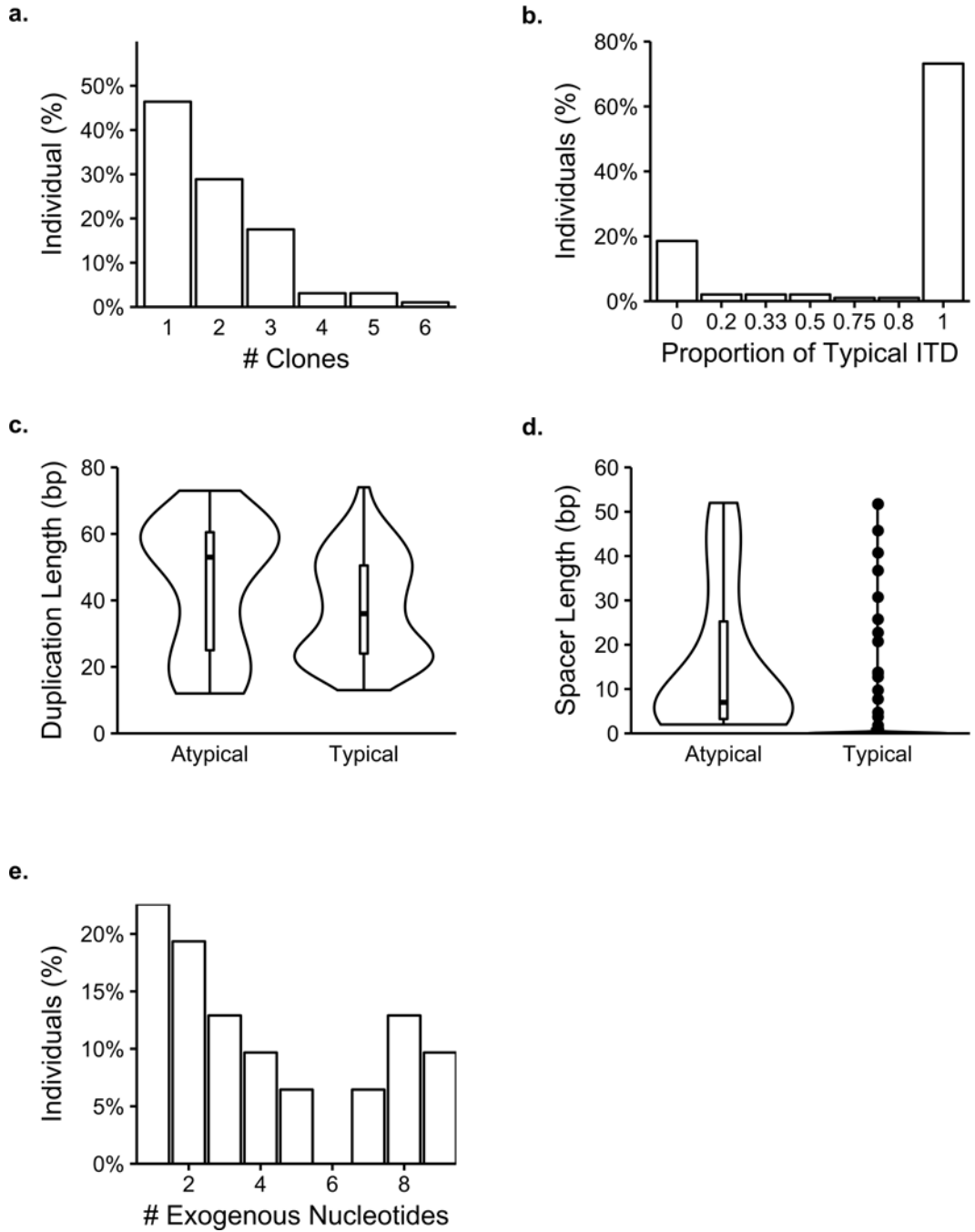
**Figure 2. Comparison of HeatITup with commonly used ITD detection algorithms.**

Comparative analysis of HeatITup with a) four commonly used methods on the capture-based WES TCGA AML cohort, where fill color represents the number of algorithms that identified a *FLT3*-ITD-positive individual; or b) a two-step CLIA-confirmed analysis workflow, consisting of a GATK/Pindel-based pipeline followed by manual variant review, on amplicon-based samples from the UPENN cohort.



**Figure 3. HeatITup-generated output of 32 detected and annotated *FLT3*-ITD mutations that were independently analyzed by a CLIA-confirmed workflow and manually annotated by expert variant reviewers.**

Light blue and green: the repeated sequence, dark blue: possible point mutations, orange: exogenous segments. Colors are customizable in the HeatITup rendered plot. Red box: CPDC140113, the manually misclassified sample. Letter suffixes of sample ID indicate different clones of a patient. Black box: magnified ITD annotation.



**Figure 4. Characteristics of typical and atypical *FLT3*-ITDs' nucleotide features.**

a) The percent of individuals per number of clones. b) The percent of individuals with different compositions of typical and atypical. The proportion of typical ITD represents ( $\#$  typical clones / ( $\#$  atypical clones +  $\#$  typical clones)). c) Violin plots of duplicated segment lengths for typical and atypical individuals (Typical median: 36 bp, Atypical median: 53, Mann-Whitney U test:  $p = 0.0750$ ). d) Violin plots of spacer segments for typical and atypical individuals (Typical median: 0, Atypical median: 7, Mann-Whitney U test:  $p <$

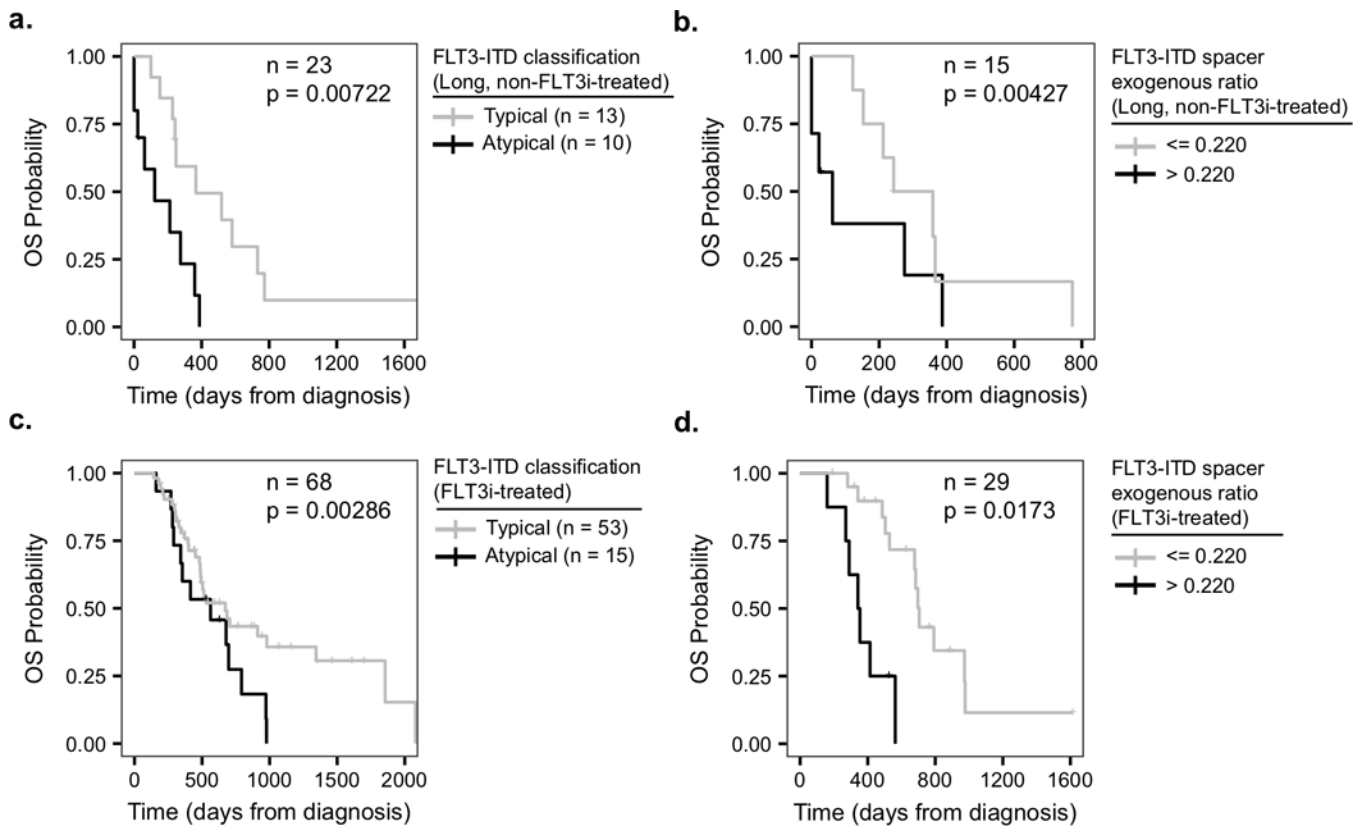
2.22e-16). e) Distribution of the number of exogenous nucleotides per individual (median (mean): 3 (4.06) bp). n = 97 individuals.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 5. Correlation between ITD features and overall survival for the *FLT3*-ITD-positive AML patients.**

a) Individuals with typical (gray, median: 366 days) vs. atypical (black, median: 122 days) clones in the non-FLT3i-treated cohort with duplication segment lengths > 29 bp (top third quartile) ( $n = 23$ , Cox PH: HR = 0.226, CI = 0.0764 – 0.669,  $p = 0.00722$ ). b) Comparison of OS between non-FLT3i-treated individuals carrying ITDs with duplication segment lengths > 29 bp (top third quartile) with various exogenous ratio values, as defined by exogenous segment length / spacer segment length ( $n = 15$ , Cox PH: HR = 159.0, CI = 4.92 – 5160,  $p = 0.00427$ ). For visualization purposes a high / low cutoff of 0.22 for the exogenous ratio, which was obtained from a maximum separation analysis, was used to plot the Kaplan-Meier curve. c) Typical (gray, median: 672 days) vs. atypical (black, median: 563 days) for all individuals in the FLT3i-treated cohort ( $n = 68$ , Cox PH: HR = 0.337, CI = 0.165 – 0.68,  $p = 0.00286$ ). d) Comparison of OS between individuals carrying ITDs with various exogenous ratio values in the FLT3i-treated cohort ( $n = 29$ , Cox PH: HR = 4.96, CI = 1.33 – 18.5,  $p = 0.0173$ ). Only for visualization purposes, the exogenous ratio of 0.22 was used to plot the Kaplan-Meier curve as in (b). Cox regressions with allogeneic transplant and disease status at the time of FLT3i treatment initiation as additional covariates were used for the FLT3i-treated cohort. In panels (b) and (d), exogenous ratio separations were only used for visualization and not Cox regressions, where the entire spectrum of exogenous ratio values were used and the corresponding p-values were reported.

**Table 1.**

Clinical characteristics of 84 FLT3-ITD-positive AML patients.

	n	%
All Subjects	84	-
Sex		
Male	43	51.2
Female	41	48.8
Age at Diagnosis		
< 60	54	64.3
60	30	35.7
Disease Status		
Newly Diagnosed <sup>A</sup>	58	69.0
Relapse	26	31.0
Induction Chemotherapy		
Yes	82	97.6
No	2	2.40
FLT3 Inhibitor Treated		
Yes	68	81.0
No	16	19.0
FLT3 Inhibitor Characterization		
Type 1	23	33.8
Type 2	29	42.7
Both	16	23.5
Allogeneic Transplant		
Yes	54	64.3
No	30	35.7
Disease Status at initiation of FLT3i		
Remission	32	38.1
Active Disease	52	61.9

<sup>A</sup>Of the 58 newly diagnosed patients, none had favorable cytogenetics, 49 had intermediate risk, 6 had adverse risk, 1 had no growth, and 2 did not have cytogenetics performed.