
Chapter 1

Genetic screening test to detect translocations in acute leukemia by use of targeted locus amplification

Clinical Chemistry 2018;64(7):1096-1103.

DOI: 10.1373/clinchem.2017.286047

PubMed ID: 29794109

M.Z. Alimohamed^{1,*}, L.F. Johansson^{1,2,*}, E.N. de Boer¹, E. Splinter³, P. Klous³, M. Yilmaz³, A. Bosga¹, M. van Min³, A.B. Mulder⁴, E. Vellenga⁵, R.J. Sinke¹, R.H. Sijmons¹, E. van den Berg¹, B. Sikkema-Raddatz¹

1. University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, The Netherlands
2. University of Groningen, University Medical Center Groningen, Genomics Coordination Center, Groningen, The Netherlands
3. Cergentis b.v., Utrecht, The Netherlands
4. University of Groningen, University Medical Center Groningen, Department of Laboratory Medicine, Groningen, The Netherlands
5. University of Groningen, University Medical Center Groningen, Department of Hematology, Groningen, The Netherlands

Received 2017 Dec 19; Accepted 2018 Apr 16; Published online May 2018.

* Contributed equally

Abstract

BACKGROUND: Over 500 translocations have been identified in acute leukemia. To detect them, most diagnostic laboratories use karyotyping, fluorescent in situ hybridization, and reverse transcription PCR. Targeted locus amplification (TLA), a technique using next-generation sequencing, now allows detection of the translocation partner of a specific gene, regardless of its chromosomal origin. We present a TLA multiplex assay as a potential first-tier screening test for detecting translocations in leukemia diagnostics. **METHODS:** The panel includes 17 genes involved in many translocations present in acute leukemias. Procedures were optimized by using a training set of cell line dilutions and 17 leukemia patient bone marrow samples and validated by using a test set of cell line dilutions and a further 19 patient bone marrow samples. Per gene, we determined if its region was involved in a translocation and, if so, the translocation partner. To balance sensitivity and specificity, we introduced a gray zone showing indeterminate translocation calls needing confirmation. We benchmarked our method against results from the 3 standard diagnostic tests. **RESULTS:** In patient samples passing QC, we achieved a concordance with benchmarking tests of 81% in the training set and 100% in the test set, after confirmation of 4 and nullification of 3gray zone calls (in total). In cell line dilutions, we detected translocations in 10% aberrant cells at several genetic loci. **CONCLUSIONS:** Multiplex TLA shows promising results as an

acute leukemia screening test. It can detect cryptic and other translocations in selected genes. Further optimization may make this assay suitable for diagnostic use.

1.1 Introduction

Molecular investigations of structural genomic aberrations and determination of the genotype have contributed to the understanding of the pathogenesis of leukemias and are essential for their diagnosis, treatment, and prognosis[25]. Currently, 500 translocations involving multiple genes have been described in hematologic malignancies, in particular, acute leukemias [7]. Routine diagnostic methods such as karyotyping, fluorescent in situ hybridization (FISH), and reverse transcription PCR (RT-PCR) are used to detect recurrent chromosomal aberrations but have limited genomic resolution or analytical sensitivity and are, at times, inadequate[22]. Translocation detection methods based on next-generation sequencing (NGS) offer several advantages over conventional clinical laboratory methods, such as the ability to detect cryptic rearrangements and unknown fusion partner genes at multiple locations simultaneously[16]. Whole-genome sequencing (WGS) can detect chromosomal translocations in acute leukemia patients[30]. However, owing to the possibly low load of leukemic cells, deep sequencing is required to reach a high sensitivity. Therefore, WGS is not yet the method of choice in a diagnostic setting. To overcome the limitations of WGS, targeted sequencing approaches can be used to analyze a specific set of genes or gene regions and detect translocation partners in cancer-related genes[6]. Despite the higher number of reads targeting genes of interest, the short length of the DNA fragments used in NGS means that only a small fraction of the reads will capture the translocation partner and be informative. One strategy to overcome this problem is to use outward orientated primers, as in the genomic inverse PCR for exploration of ligated breakpoints (GIPFEL) technique, which can detect chromosomal translocations in childhood leukemia[10]. However, targeted methods such as GIPFEL require prior knowledge of both the translocation partners and the genomic locations of breakpoints. The many possible breakpoints and gene fusion partners limit the applicability of such techniques and make these techniques less suitable as stand-alone techniques in a diagnostic setting. A more robust, comprehensive, and unbiased method for detection of translocations is therefore required. A recently reported technique, targeted locus amplification (TLA), enables translocations to be detected regardless of the identity of the chromosomal partner[5]. TLA uses the principles of proximity ligation of crosslinked DNA, followed by targeted amplification using

outward-orientated primers and subsequent sequencing of any locus of interest, thereby capturing hundreds of kilobases of surrounding DNA[5]. TLA can thus detect translocations involving a gene of interest without prior knowledge of the fusion partner and allows the breakpoint to be located at some distance from the probe used, potentially capturing novel translocation partners. We aimed to develop a TLA assay as a first-tier screening test to detect translocations in acute leukemia. Here we present a comprehensive, multiplex gene panel designed to cover 17 common genes involved in acute leukemias and known to be associated with hundreds of fusion gene partners. In this proof-of-principle study, we compared the clinical utility of targeted translocation detection using our acute leukemia NGS gene panel with the results from current genetic diagnostic tests in a series of patient bone marrow samples.

1.2 Material and Methods

1.2.1 Patient bone marrow cells and cell lines

Bone marrow cells were obtained from 36 patients diagnosed with leukemia following informed consent. The study protocol was approved by the Ethics Committee of the University Medical Centre Groningen (METC 2014.051, 10–2-2014). The cells were washed with 1X red blood cell lysis buffer (Stem Cell Technologies). Mononuclear cells were isolated by centrifugation (10 min at 250g) and stored in complete RPMI 1640 culture medium (Lonza), supplemented with 10% v/v DMSO (Merck KGaA) at -140 °C. In addition, we used 5 different cell lines, carrying known translocations that included genes present in our panel: KOPN-8 [t(11;19)(q23;p13), t(8;13)(q24;q21.1); lysine methyltransferase 2A (*KMT2A*), MYC proto-oncogene, bHLH transcription factor (*MYC*)]; HAL-01 [t(17;19)(q22;p13); transcription factor 3 (*TCF3*)]; FKH-1 [t(6;9)(p23;q34); DEK proto-oncogene (*DEK*)]; REH [t(12;21)(p13;q22); ETS variant 6 (*ETV6*)/runt related transcription factor 1 (*RUNX1*)]; and MV4-11[t(4;11)(q21;q23); *KMT2A*; all from Leibniz Institute DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen); see Table 1 in the Data Supplement that accompanies the online version of this article at <http://www.clinchem.org/content/vol64/issue7>]. Cell line GM12878 (Coriell institute) was used to test the multiplex primer quality (see Fig. 1 in the online Data Supplement). All cell lines used were cultured according to the instructions provided by their repository.

1.2.2 TLA acute leukemia gene panel

Seventeen genes involved in gene fusions associated with acute leukemia were selected [ABL proto-oncogene 1, non-receptortyrosinekinase(*ABL1*), baculoviral IAP repeat containing 3 (*BIRC3*), core-binding factor subunit beta (*CBFB*), *DEK*, *ETV6*, fibroblast growth factor receptor 1 (*FGFR1*), homeobox A9 (*HOXA9*), lysine acetyltransferase 6A (*KAT6A*), *KMT2A* (*MLL*), *MYC*, nucleophosmin1 (*NPM1*), phosphatidylinositol binding clathrin assembly protein (*PICALM*), retinoic acid receptor alpha (*RARA*), RNA binding motif protein 15 (*RBM15*), *RUNX1*, *TCF3*, and zinc finger MYM-type containing 2 (*ZMYM2*)]. Target regions within these genes are involved in numerous chromosomal translocations and were defined according to known breakpoints reported in the literature[25, 7, 16, 6, 2] (see Table 2 in the online Data Supplement). To enable comprehensive coverage of the target regions, we designed 43 inverse primer sets. After single primer testing, the primers were placed in optimal concentrations in 2 multiplex assays. Multiplex 1 consisted of 26 primer sets designed to cover known breakpoint regions, whereas multiplex 2 had 17 primer sets to boost the coverage around the target regions (see Table 3 and Fig.2 in the online Data Supplement).

1.2.3 Multiplex TLA sample preparation, sequencing, and data analysis

Before TLA was performed, cells were harvested (cell lines) or thawed (bone marrow cells) and washed with RPMI 1640 media, and the concentration was determined using the average of 3 cell counts (Sysmex KX21N; Sysmex Corporation). A total of $5 - 10 \times 10^6$ cells were used as starting material. Cell suspensions were homogenized and TLA was performed separately for multiplexes 1 and 2 according to the manufacturer's protocol[5]. Full protocols are described in the online Data Supplement. In short, purified circular DNA fragments were sheared, end-repaired, dA-tailed, and adapter-ligated. Fragments in the 300- to 320-bp range were equimolarly pooled per 24 samples and loaded at a concentration of 0.65 pmol/L on a NextSeq 500 platform (Illumina) using a high-output flow cell kit having paired end reads at 2×151 -bp read length and v2 reagents. Using a set of training samples as described below, we set up the data analysis procedure. In short, duplicate reads were removed and digested in silico at CATG sites (the *NotI* restriction site used in the TLA procedure). Reads were aligned to the human genome (build 37) and split into 17 separate files, 1 for each region of interest. Then, for each region of interest, reads were counted in 10-kb bins and filtered. We determined the presence of peaks and represented them on genome-wide plots

and in a tabulated report. Data QC was assessed and standardized based on peak width and noise level, leading to a quality label for each sample and region of interest. Based on the quality of the sample and region, and on the size of the captured peak on the potential translocation partner, we made a definitive translocation call or a gray zone translocation call needing confirmation (see Methods—Translocation calling in the online Data Supplement). This was generalized for all targets.

1.2.4 Routine genetic and cytogenetic methods

Karyotyping and additional FISH were performed according to Dutch national guidelines[26]. We analyzed a total of 20 GPG-banded metaphase cells in all patient samples and cell lines with karyotyping. FISH was performed on BCR, RhoGEF, and GTPase activating protein (*BCR*)/*ABL1*, *ETV6*, *ETV6*/*RUNX1*, MDS1 and EVI1 complex locus (*MECOM*), *FGFR1*, *KMT2A*, *MYC*, promyelocytic leukemia (*PML*)/*RARA*, *RUNX1*/*RUNX1* translocation partner 1 (*RUNX1T1*) or T cell leukemia homeobox 3 (*TLX3*)-*NPM1* or in samples having an inconclusive karyotype (see Table 4 in the online Data Supplement). The KOPN-8 cell line was also analyzed using an *MYC* breakpoint probe (KBI-10611; Kreatech) to confirm the presence of a breakpoint in or near *MYC* (see Fig. 3 in the online Data Supplement). In addition, we isolated RNA from mononuclear cells in the bone marrow, and performed reverse transcription to prepare cDNA. RT-PCR using fusion-gene specific primers was performed according to the methods used by van Dongen et al.[29] to detect the most frequent chromosomal rearrangements of leukemia: *BCR-ABL*, *ETV6-RUNX1*, *PML-RARA*, *RUNX1-RUNX1T1*.

1.2.5 Validation of the multiplex TLA method

Samples were processed in 2 sets: (a) a training set used to optimize analysis and interpretation procedure and (b) a test set used to validate the procedures.

Training set. The training set consisted of 17 patient samples with a karyotype known to the researcher and the REH and FKH-1 cell lines, as well as a cell line dilution series using the KOPN-8 and HAL-01 cell lines (see Table 5.1 in the online Data Supplement). All samples were used to set filter thresholds for data analysis and interpretation. The dilution series were also used to determine the minimum percentage of aberrant cells detectable at our set thresholds.

Test set. To assess the performance of the multiplex TLA procedure, we selected, anonymized, and tested a set of 19 patient bone marrow samples,

as described above, using the optimum thresholds from the training set analysis. In the test set we repeated the dilution series using mixed cell lines of KOPN-8, HAL-01, FKH-1, and MV4–11 (see Table 5.2 in the online Data Supplement) to confirm the minimum percentage of aberrant cells detectable by our assay. Sensitivity was further assessed by random downsampling of aligned reads of the test set’s cell line dilution series (see Table 6 in the online Data Supplement). The outcomes were benchmarked against the results obtained from routine genetic tests. A finding was considered true-positive if it was concordant in the TLA and routine diagnostic tests; it was considered true-negative if it was not detected by any of the tests. A sample finding was considered false-negative if the translocation involving genes present in the multiplex TLA panel was not detected by the TLA assay but was seen in routine tests. It was considered false-positive if the TLA assay indicated the presence of a translocation, but the routine tests could not detect it.

1.3 Results

1.3.1 Validation of the TLA multiplex panel - Training set

Optimized analysis and interpretation of patient bone marrow samples. We optimized translocation calling by the TLA multiplex pipeline by adding data filtering and defining data QC and data interpretation steps according to the location and size of the captured peaks (see Methods in the online Data Supplement). Using the analysis settings optimized for the 17 bone marrow samples, 16 samples, including 88% of targets, passed our QC. Sample #13 failed because of the absence of sequence reads (peaks) on target regions owing to a low cell count (see Table 7 in the online Data Supplement) and was eliminated from further analysis. In total, 9 definitive translocation calls were made (Table1). In sample #9 there were 2 separate events (see Table 8 in the online Data Supplement). The first involved captured peaks smaller than the threshold in the ABL1 and MYC targets, resulting in a translocation call in the gray zone that required confirmation. Procedure-wise, this call was followed up, which led to further evaluation using the karyotype information, after which the gray zone call was considered negative. A translocation known to be present in sample #9, t(8;21)(q22;q22), led to a false-negative result, because the expected peak on chromosome 8, from the RUNX1 viewpoint, was not detected. In a further 2 samples, a translocation was missed. These translocations were labeled as false negatives. In 1 of these samples [#10, t(11; 19)(q23;p13.1)], as well as the earlier mentioned sample #9 – t(8;21)(q22;q22), multiplex amplification on the targeted region was not able to generate a sufficient number of reads on the translocation partner to

pass our data filter threshold. In sample #5, t(11;17)(q23;q25) was missed. Here, there were no reads present on the translocation partner. Four other samples in which the multiplex TLA panel detected no translocations had normal karyotypes. No false-positive results were found. In total, 13 of the 16 samples passing the QC generated concordant results to routine genetic and cytogenetic results (see Table 9 and Fig. 5 in the online Data Supplement).

Sensitivity to detect translocations present in a low percentage of cells. Dilution series of the cell lines KOPN-8 and HAL-01 with 5% to 100% aberrant cells were prepared to determine the translocation detection sensitivity of the TLA panel. Optimized analysis settings using the filter and interpretation steps labeled all samples and 94% of targeted regions as passing QC. In the HAL-01 cell line, t(17;19)(q22;p13), including *TCF3*, was seen in samples having at least 10% aberrant cells. This also holds for t(11;19)(q23;p13) in the KOPN-8 cell line, including *KMT2A*. In the same cell line, t(8;13)(q24;q21.2) was seen in the presence of 25% aberrant cells. In total, above 10% aberrant cells, 20 translocation calls were made, of which 3 were labeled as gray-zone. All calls were positive after confirmation. *MYC* was not detected at 10% aberrant cells, leading to detection of 20 out of 21 translocations (see Table 8 in the online Data Supplement). No false-positive calls were made in the cell line training set. As additional positive controls for complex cryptic translocations, the cell lines REH and FKH-1 were tested on samples with 100% aberrant cells. Cell line REH was previously karyotyped (see Table 1 in the online Data Supplement) as carrying a 4-way translocation t(4;12;21;16)(q32;p13;q22;q24.3)[19]. TLA did not find partner chromosome 4 from the position of the chromosome 12 target region, although it successfully detected partner chromosome 21. TLA also detected chromosomal partners 12 and 16 captured from the target region on chromosome 21 (see Table 8 in the online Data Supplement). TLA results were confirmed by additional karyotyping, leading to recharacterization of the translocation to t(12;21;16)(p13;q22;q24.3). In cell line FKH-1, we detected t(6;9)(p23;q34), resulting in a DEK-nucleoporin 214 (*NUP214*) fusion gene. In addition, we identified a translocation t(9;12)(q34;p13), which was not present in the cytogenetic information of the cell line catalogue[18]. FISH using the *BCR/ABL1* and *ETV6* probes supports this finding (see Fig. 4 in the online Data Supplement).

1.3.2 Validation of the TLA multiplex panel - Test set

High concordance between TLA multiplex panel and routine tests for bone marrow samples. We assessed the clinical utility of the optimized and fixed data analysis and interpretation procedure on anonymized test set samples.

Results from the TLA procedure and routine genetic tests were compared. A total of 14 out of 19 samples, including 74% of targets, passed QC. All 5 samples that failed QC (#25, 26, 28, 29, and 33) were from nonhomogeneous cell suspensions, leading to the absence of sequence reads (peaks) on target regions. In the samples that passed QC, we made 3 definitive translocation calls and 6 gray zone calls needing confirmation. Four of the 6 gray zone calls, 3 t(12;21) and 1 t(3;12), were confirmed (Table 1.1) by other genetic tests. Two were considered negative after follow up with confirmatory tests and routine diagnostic data. We detected no translocations in 7 samples. All translocation calls were concordant with the benchmarking tests. No translocations involving genes present in the multiplex TLA panel were missed and no false-positive translocations were called (see Table 9 and Fig. 5 in the online Data Supplement).

Reproducibility of sensitivity to detect translocations in aberrant cell lines. We performed a second dilution series in a range of 1% to 50% aberrant cells, involving test cell lines (KOPN-8, HAL-01, FKH-1, and MV4-11) to confirm the translocation detection sensitivity of the TLA panel in repeated cell line samples. Translocations including *TCF3*, *DEK*, *ETV6*, *RUNX1*, and *KMT2A* were detected in samples with a minimum of 10% aberrant cells. Similar to the training set dilution series, all the test set samples, including 99% of targets, passed QC. The translocation involving *MYC* was detected in samples containing at least 25% aberrant cells. In total, above 10% aberrant cells, 18 translocation calls were made, of which 4 were labeled as gray zone. After confirmation, 2 of the gray zone calls were positive and 2 were nullified. *MYC* was not detected at 10% aberrant cells, leading to detection of 16 out of 17 translocations—including FKH-1 t(9;12)(q34;p13) (see Table 8 in the online Data Supplement). No false positive translocation calls were made.

CHAPTER 1. TLA

Table 1.1: TLA and benchmarking of the results from the training and test sets

	Sample	Referral reason	Translocation in ROI	Karyotype	FISH	RT- PCR	TLA
1	Training set						
	1	ALL	t(12;21)(p13;q22)	1 ¹	n/a	+ ²	+
2	3	ALL	t(8;14)(q24;q32)	+	n/a	n/a	+
	4	CML	t(9;22)(q34;q11.2)	+	n/a	+	+
3	5	AML	t(11;17)(q23;q25)	+	+* ³	n/a	-
	6	CML	t(9;22)(q34;q11.2)	+	n/a	+	+
	9	AML	t(8;21)(q22;q22)	+	+	+	-
4	10	AML	t(11;19)(q23;p13.1)	+	+*	n/a	-
	11	AML	t(9;22)(q34;q11.2)	+	+	+	+
	12	ALL	t(1;19)(q23;p13.3)	+	n/a	n/a	+
5	13	AML	t(15;17)(q34;q11.2)	+	+	+	n/a
	15	CML	t(9;22)(q34;q11.2)	+	+	+	+
6	16	AML	t(15;17)(q24;q21)	+	+	+	+
	17	ALL	t(4;11)(q21;q23)	+	+*	n/a	+
	2	ALL	None	-	-	-	-
7	7	AML	None	-	-	n/a	-
	8	AML	None	-	n/a	-	-
8	14	AML	None	-	-	n/a	-
	Test set						
9	18	AML	t(9;22)(q34;q11.2)	+	+	+	+
	19	AML	t(11;19)(q23;p13.1)	+	+*	n/a	+
	20	AML	t(3;12)(q26;p12)	+	+*	n/a	+
10	24	ALL	t(12;21)(p13;q22)	-	n/a	+	+
	26	AML	t(9;22)(q34;q11.2)	+	n/a	+	+
11	30	ALL	t(12;21)(p13;q22)	-	n/a	+	+
	35	ALL	t(12;21)(p13;q22)	-	n/a	+	+
	36	ALL	t(12;21)(p13;q22)	-	+	+	+
	21	ALL	None	-	-	n/a	-
	22	AML	None	-	n/a	n/a	-
	23	AML	None	-	-	-	-
	25	ALL	None	-	-	n/a	n/a
	27	ALL	None	-	-	n/a	-
	28	AML	None	-	-	-	n/a
	29	ALL	None	-	-	n/a	n/a
	31	ALL	None	-	-	-	-
	32	ALL	None	-	-	-	-
	33	ALL	None	-	-	-	n/a
	34	ALL	None	-	-	-	-

[1] (-) Translocation absent

[2] (+) Translocation present

[3] (+*) Break seen on 1 translocation partner

1.4 Discussion

We have developed a genetic screening assay to detect translocations relevant to acute leukemia using a multiplex TLA panel in combination with NGS. The TLA assay allows screening of multiple genomic regions and numerous samples simultaneously on a single platform, including those with cryptic translocations such as t(12; 21). Up to now, karyotyping, in combination with FISH and/or RT-PCR, has been required to detect such translocations [22]. Using our assay, we were able to detect translocations in cell lines with at least 10% aberrant cells for the genes tested (*MYC* at 25%). This sensitivity is in the same range as that offered by karyotyping [20, 4], although karyotyping often fails in detecting cryptic translocations and complex aberrations, and it also needs cells to be cultured. RT-PCR and FISH have sensitivities of 0.01%–1% [3, 21] and 5%–10% [23, 1, 14], respectively, but these tests only work on specific targeted translocations or give no information on the translocation partner. Our TLA assay offers a competitive option for screening of unknown and cryptic translocation partners. For the patient samples that passed QC, we achieved a concordance with routine genetic testing of 81% in the training set and 100% in the test set for detecting translocations involving genes included in our TLA multiplex panel. In the training set, 2 translocations, t(8;21)(q22; q22) and t(11;19)(q23;p13), had too few reads to be distinguished from background signal. It is likely that the 5 million cells used in the assay were suboptimal in yielding sufficient quality for the detection of rearrangements. This was solved by doubling the number of cells used, which led to detection of all targeted translocations in test samples. We have observed that some targets are susceptible to suboptimal sample quality, resulting in inadequate enrichment. We also found that a nonhomogeneous cell suspension and clots in frozen samples yielded low-quality results. We therefore recommend starting with fresh material or assessing cell viability after thawing of frozen samples and starting the TLA procedure with 10 million viable cells. Primer concentrations for sensitive targets such as *BIRC3*, *CBFB*, and *KAT6A* need further optimization to improve the robustness of the panel. The third translocation we missed was t(11;17) (q23;q25). This translocation was present in around 70% of karyotyped metaphases of sample #5. However, no reads were present on chromosome 17 in the TLA multiplex panel, although FISH demonstrated the chromosome 11 breakpoint to be in the *KMT2A* gene. Often, seemingly balanced translocations are accompanied by deletions[16]. Both *KMT2A* probes used in our panel are located within 10 kb of the major breakpoint region between exons 7 and 13[3]. A possible explanation for the missed translocation would be a deletion of this region. This will result in the absence of *KMT2A* probe target locations on

the translocation chromosome and a subsequent false-negative result. Such known complexity around breakpoints should be taken into account when designing the panel and interpreting the results. Including an additional TLA primer set further from the expected breakpoint could avoid this problem in future experiments. In our sample cohort we obtained 100% specificity with no false-positive results. Another multiplex TLA panel, described earlier, targeting 19 BCP-ALL (B-cell precursor acute lymphoblastic leukemia) genes, identified all known rearrangements but did not mention the specificity of the panel[12]. Occasionally, owing to a low percentage of aberrant cells in a sample, a translocation can be missed when using strict analysis thresholds. To ensure optimal sensitivity and specificity, we introduced a gray zone. Captured chromosomal regions with few reads suggesting the presence of a translocation are considered as a gray zone translocation call. This enables a more explicit assessment of indeterminate translocation calls that have a moderate coverage to prevent false positive calls and missed translocations. In such cases, an additional confirmation test is required. In our small test cohort, starting with an optimal number of cells, 67% of the gray zone calls were confirmed. Using this strategy, a 100% sensitivity and specificity was obtained. Our multiplex TLA assay potentially captures all translocations involving 1 of the 17 targeted genes up to breakpoint distances of several hundreds of kilobases. It is illustrative that when TLA was applied to samples containing complex structural variation, it resulted in the recharacterization of the genotypes described earlier in cell lines REH and FKH-1. Our screening assay identified translocations in not only targeted genes, but also in genes not directly targeted, such as *NUP214*, located 200 kb from *ABL1*, which led to detecting the *DEK-NUP214* fusion in the FKH-1 cell line, even with only 10% aberrant cells present. Furthermore, we showed that the TLA panel can be applied in other hematological malignancies, because it can detect the t(9;22)(q34;q11) and t(8; 14)(q24;q32) translocations that are found in up to 95% of patients with chronic myeloid leukemia and in 80% of those with Burkitt lymphoma [8, 9]. Using the panel, we detected 3 different *KMT2A* translocations in our small cohort: 2, t(11;19)(q23; p13.3)[*KMT2A-ENL*] and t(4;11)(q21;q23) [*KMT2A-AF4*], in the cell lines, and 1, t(11;19)(q23;p13.1), in a test set sample, likely resulting in a *KMT2A-ELL* gene fusion (Table 1.1). The *KMT2A* gene alone has 80 known fusion partners [17, 31]. Likewise, the other 16 panel genes can, in principle, detect all known translocations as well as novel ones. In contrast, other methods such as translocation comparative genomic hybridization[11] and GIPFEL[10] detect only specific fusions and show lower sensitivity. Alternatively, RNA-based techniques could be considered for translocation detection [15, 27, 24, 13, 32] and are major competitors to the TLA assay. RNA-based platforms are instrumental in the detection of

single-nucleotide variants, insertions, deletions, copy number changes, and fusions[32]. However, these techniques can detect only translocations with breakpoints in exonic or intronic regions and are dependent on the expression of the fusion gene, limiting their use for the detection of non-transcript altering translocations such as those involving MYC. We determined that our assay required a minimum of 10% aberrant cells in a sample to detect translocations involving targeted regions, with the exception of the MYC target region, where the detection limit was 25% for t(8;13)(q24;q21.1). A likely explanation for this lower sensitivity is that, for MYC, probes were designed solely in the 190-kb region associated with the most common breakpoint regions for translocations t(8;14), t(2;8), and t(8;22), whereas it is known that breakpoints around MYC can be present in a much larger area of 2 Mb [28]. However, we reduced the location of the breakpoint to the 740-kb region covered by the MYC break-apart probe, of which 620 kb is located distally from our targeted breakpoint region. It is therefore possible that the breakpoint of the rare t(8;13) translocation is located outside our region of interest, making it harder to capture the translocation partner and thus lowering the sensitivity. In conclusion, in this proof-of-principle study, our multiplex TLA assay shows promising results that indicate it is suitable as a first-tier screening test in acute leukemia, chronic myeloid leukemia, and Burkitt lymphoma for detection of most common cryptic and other translocations, without prior knowledge of particular fusion partners. Further improvements in probe concentrations, input quality control, and automation of total workflow will enhance robustness and sensitivity and may make the assay suitable for diagnostic use.

1.5 Acknowledgments

We thank Jackie Senior and Kate Mc Intyre for editorial advice.

Authors' Disclosures or Potential Conflicts of Interest

Disclosures and/or potential conflicts of interest: Employment or Leadership: E. Splinter, Cergentis b.v.; P. Klous, Cergentis b.v.; M. Yilmaz, Cergentis b.v.; M. van Min, Cergentis b.v. Consultant or Advisory Role: None declared. Stock Ownership: M. van Min, Cergentis b.v. Honoraria: None declared. Research Funding: ZONMW, grant no 40-41200-98-9159. Expert Testimony: None declared. Patents: None declared. Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or final approval of manuscript

1.6 Online data supplement

<http://clinchem.aaccjnls.org/content/clinchem/suppl/2018/04/27/clinchem.2017.286047.DC1/clinchem.2017.286047-1.pdf>

1

2

3

4

5

6

7

8

9

10

11

Bibliography

- [1] Umut Aypar, Ryan A. Knudson, Kathryn E. Pearce, Anne E. Wiktor, and Rhett P. Ketterling. Development of an npml1/mlf1 d-fish probe set for the detection of t(3;5)(q25;q35) identified in patients with acute myeloid leukemia. *The Journal of Molecular Diagnostics*, 16(5):527–532, Sep 2014.
- [2] S.K. Bohlander. Fusion genes in leukemia: an emerging network. *Cytogenetic and Genome Research*, 91(1-4):52–56, 2000.
- [3] Thomas Burmeister, Claus Meyer, Daniela Gröger, Julia Hofmann, and Rolf Marschalek. Evidence-based rt-pcr methods for the detection of the 8 most common mll aberrations in acute leukemias. *Leukemia Research*, 39(2):242–247, Feb 2015.
- [4] Jeremy J Buzzard, Nicholas M Gough, Jeremy M Crook, and Alan Colman. Karyotype of human es cells during extended culture. *Nature Biotechnology*, 22(4):381–382, Apr 2004.
- [5] Paula J P de Vree, Elzo de Wit, Mehmet Yilmaz, Monique van de Heijning, Petra Klous, Marjon J A M Verstegen, Yi Wan, Hans Teunissen, Peter H L Krijger, Geert Geeven, and et al. Targeted sequencing by proximity ligation for comprehensive variant detection and local haplotyping. *Nature e Biotechnology*, 32(10):1019–1025, Aug 2014.
- [6] Eric J Duncavage, Haley J Abel, Philippe Szankasi, Todd W Kelley, and John D Pfeifer. Targeted next generation sequencing of clinically significant gene mutations and translocations in leukemia. *Modern Pathology*, 25(6):795–804, Mar 2012.
- [7] Mitelman F., Johansson B., and Mertens F (Eds.). Mitelman database of chromosome aberrations and gene fusions in cancer 2017, 2017.
- [8] Stefan Faderl, Moshe Talpaz, Zeev Estrov, Susan O'Brien, Razelle Kurzrock, and Hagop M. Kantarjian. The biology of chronic myeloid leukemia. *New England Journal of Medicine*, 341(3):164–172, Jul 1999.

BIBLIOGRAPHY

- [9] J. A. Ferry. Burkitt's lymphoma: Clinicopathologic features and differential diagnosis. *The Oncologist*, 11(4):375–383, Apr 2006.
- [10] Elisa Fueller, Daniel Schaefer, Ute Fischer, Pina F. I. Krell, Martin Stanulla, Arndt Borkhardt, and Robert K. Slany. Genomic inverse pcr for exploration of ligated breakpoints (gipfel), a new method to detect translocations in leukemia. *PLoS ONE*, 9(8):e104419, Aug 2014.
- [11] Harvey A. Greisman, Noah G. Hoffman, and Hye Son Yi. Rapid high-resolution mapping of balanced chromosomal rearrangements on tiling cgk arrays. *The Journal of Molecular Diagnostics*, 13(6):621–633, Nov 2011.
- [12] R.P. Kuiper, S.V. van Reijmersdal, M. Simonis, J. Yu, E. Sonneveld, B. Scheijen, and et al. Targeted locus amplification and next generation sequencing for the detection of recurrent and novel gene fusions for improved treatment decisions in pediatric acute lymphoblastic leukemia. *Blood*, 126(23):696, 2015.
- [13] Joshua Z Levin, Michael F Berger, Xian Adiconis, Peter Rogov, Alexandre Melnikov, Timothy Fennell, Chad Nusbaum, Levi A Garraway, and Andreas Gnirke. Targeted next-generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts. *Genome Biology*, 10(10):R115, 2009.
- [14] Michael Liew, Leslie Rowe, ParkerW Clement, RodneyR Miles, and MohamedE Salama. Validation of break-apart and fusion myc probes using a digital fluorescence in situ hybridization capture and imaging system. *Journal of Pathology Informatics*, 7(1):20, 2016.
- [15] H Lilljebjörn, H Ågerstam, C Orsmark-Pietras, M Rissler, H Ehrencrona, L Nilsson, J Richter, and T Fioretos. Rna-seq identifies clinically relevant fusion genes in leukemia including a novel mef2d/csf1r fusion responsive to imatinib. *Leukemia*, 28(4):977–979, Nov 2013.
- [16] Fredrik Mertens, Bertil Johansson, Thoas Fioretos, and Felix Mitelman. The emerging complexity of gene fusions in cancer. *Nature Reviews Cancer*, 15(6):371–381, Jun 2015.
- [17] C Meyer, J Hofmann, T Burmeister, D Gröger, T S Park, M Emerenciano, M Pombo de Oliveira, A Renneville, P Villarese, E Macintyre, and et al. The mll recombinome of acute leukemias in 2013. *Leukemia*, 27(11):2165–2176, Apr 2013.
- [18] Leibniz Institut DSMZ-German Collection of Microorganisms and Cell Cultures. Fkh1:dsmz no acc 614, 2017.
- [19] Leibniz Institut DSMZ-German Collection of Microorganisms and Cell Cultures. Reh:dsmz no acc22, 2017.
- [20] D. Pinkel, T. Straume, and J.W. Gray. Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc Natl Acad Sci USA*, 83:2934–8, 1986.
- [21] Manuel Salto-Tellez, Suresh G. Shelat, Bernice Benoit, Hanna Rennert, Martin Carroll, Debra G.B. Leonard, Peter Nowell, and Adam Bagg. Multiplex rt-pcr for the detection of leukemia-associated translocations. *The Journal of Molecular Diagnostics*, 5(4):231–236, Nov 2003.
- [22] Avery A. Sandberg and Aurelia M. Meloni-Ehrig. Cytogenetics and genetics of human cancer: methods and accomplishments. *Cancer Genetics and Cytogenetics*, 203(2):102–126, Dec 2010.

-
- [23] C Schoch, S Schnittger, S Bursch, D Gerstner, A Hochhaus, U Berger, R Hehlmann, W Hiddemann, and T Haferlach. Comparison of chromosome banding analysis, interphase- and hypermetaphase-fish, qualitative and quantitative pcr for diagnosis and for follow-up in chronic myeloid leukemia: a study on 350 cases. *Leukemia*, 16(1):53–59, Jan 2002.
- [24] Jonathan A. Scolnick, Michelle Dimon, I-Ching Wang, Stephanie C. Huelga, and Douglas A. Amorese. An efficient method for identifying gene fusions by targeted rna sequencing from fresh frozen and ffpe samples. *PLOS ONE*, 10(7):e0128916, Jul 2015.
- [25] Lisa G Shaffer, Roger A Schultz, and Blake C Ballif. The use of new technologies in the detection of balanced translocations in hematologic disorders. *Current Opinion in Genetics and Development*, 22(3):264–271, Jun 2012.
- [26] S. Snijder, B. Beverloo, C. Mellink, M. Stevens-Kroef, E. van den Berg, Buijs A., and et al. Vkg1 v07: Richtlijnen verworven cytogenetica., 2015.
- [27] Synne Torkildsen, Ludmila Gorunova, Klaus Beiske, Geir E. Tjønnfjord, Sverre Heim, and Ioannis Panagopoulos. Novel zeb2-bcl11b fusion gene identified by rna-sequencing in acute myeloid leukemia with t(2;14)(q22;q32). *PLOS ONE*, 10(7):e0132736, Jul 2015.
- [28] E. Van den Berg and Stevens-Kroef M. t(8;14)(q24;q32) igh/myc; t(2;8)(p12;q24) igk/myc; t(8;22)(q24;q11) igl/myc. atlas genet cytogenet oncol heamatol, 2017.
- [29] J.J. van Dongen, E.A. MacIntyre, E. Delabesse, V. Rossi, G. Saglio, and et al. Standardized rt-pcr analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease. report of the biomed-1 concerted action: investigation of minimal residual disease in acute leukemia.
- [30] John S. Welch. Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. *JAMA*, 305(15):1577, Apr 2011.
- [31] Amanda C. Winters and Kathrin M. Bernt. Mll-rearranged leukemias—an update on science and clinical approaches. *Frontiers in Pediatrics*, 5, Feb 2017.
- [32] Zongli Zheng, Matthew Liebers, Boryana Zhelyazkova, Yi Cao, Divya Panditi, Kerry D Lynch, Juxiang Chen, Hayley E Robinson, Hyo Sup Shim, Juliann Chmielecki, and et al. Anchored multiplex pcr for targeted next-generation sequencing. *Nature Medicine*, 20(12):1479–1484, Nov 2014.

List of Tables

1.1	TLA and benchmarking of the results from the training and test sets	12
-----	---------------------------------------------------------------------	----

List of Figures