Published in final edited form as:

Eur J Cancer. 2014 March ; 50(4): 793–800. doi:10.1016/j.ejca.2013.11.030.

Rapid identification of compound mutations in patients with Ph-positive leukemias by long-range next generation sequencing

R. Kastner^{1,2}, A. Zopf³, S. Preuner^{1,2}, J. Pröll³, N. Niklas³, P. Foskett⁴, P. Valent⁵, T. Lion^{*},
1,2,6, and C. Gabriel³

¹Children's Cancer Research Institute, Vienna, Austria

²Labdia Labordiagnostik GmbH, Vienna, Austria

³Red Cross Transfusion Service for Upper Austria, Linz, Austria

⁴Imperial Molecular Pathology Laboratory, Hammersmith Hospital, Imperial College Healthcare NHS Trust, London, UK

⁵Department of Medicine I, Division of Hematology and Hemostaseology and Ludwig Boltzmann Cluster Oncology, Medical University of Vienna, Vienna, Austria

⁶Department of Pediatrics, Medical University of Vienna, Austria

Abstract

An emerging problem in patients with Ph-positive leukemias is the occurrence of cells with multiple mutations in the *BCR-ABL1* tyrosine kinase domain (TKD) associated with high resistance to different tyrosine kinase inhibitors. Rapid and sensitive detection of leukemic subclones carrying such changes, referred to as compound mutations, is therefore of increasing clinical relevance. However, current diagnostic methods including next generation sequencing (NGS) of short fragments do not optimally meet these requirements. We have therefore established a long-range (LR) NGS approach permitting massively parallel sequencing of the entire TKD length of 933bp in a single read using 454 sequencing with the GS FLX+ instrument (454 Life Sciences). By testing a series of individual and consecutive specimens derived from six patients with chronic myeloid leukemia, we demonstrate that long-range NGS analysis permits sensitive identification of mutations and their assignment to the same or to separate subclones. This approach also facilitates readily interpretable documentation of insertions and deletions in the entire BCR-ABL1 TKD. The long-range NGS findings were reevaluated by an independent technical approach in select cases. PCR amplicons of the BCR-ABL1 TKD derived from individual specimens were subcloned into pGEM®-T plasmids, and >100 individual clones were subjected to analysis by Sanger sequencing. The NGS results were confirmed, thus documenting the reliability of the new technology. Long-range NGS analysis therefore provides an economic approach to the identification of compound mutations and other genetic alterations in the entire

^{*}Corresponding author and address for correspondence: Prof. Thomas Lion, MD, PhD, Children's Cancer Research Institute, Zimmermannplatz 8, A-1090 Vienna, Austria, thomas.lion@ccri.at.

Conflicts of interest statement

K.R., Z.A., P.S., P.J., N.N., F.P., and G.C. declare no conflicts of interest. V.P. received research funding and honoraria from Novartis. L.T. received research funding and honoraria from Novartis, Bristol-Myers Squibb and Pfizer.

BCR-ABL1 TKD, and represents an important advancement of the diagnostic armamentarium for rapid assessment of impending resistant disease.

Keywords

NGS; CML; ALL; *BCR-ABL1*; tyrosine kinase domain; TKI

Introduction

Mutations in the BCR-ABL1 tyrosine kinase domain (TKD) are regarded as the most important mechanism of resistance to tyrosine kinase inhibitors (TKIs) in patients with Ph-positive leukemias (1–4). A variety of methods are currently used for the detection of mutations in the BCR-ABL1 TKD ranging from broad screening techniques to highly sensitive detection methods for specific mutations (5–13). The heterogeneity of diagnostic approaches accounts, at least in part, for the differences in reported frequencies of mutations (3). Mutational screening is most commonly performed by bidirectional Sanger sequencing of the entire BCR-ABL1 TKD amplified by PCR, which does not reveal the presence of mutant subclones representing less than 10–20% of the Ph-positive cell pool (14). Despite this limitation, Sanger sequencing is currently recommended as the method of choice by the European Leukemia Net (ELN) (14–17). A number of alternative methodologies providing higher sensitivity have been introduced to facilitate earlier detection of mutant, potentially resistant BCR-ABL1 cells, and some of these approaches permit accurate monitoring of the size of mutant subclones during therapy (6, 9, 11, 13, 18–21). Our observations based on a quantitative ligation-dependent (LD) PCR technique revealed that mutant subclones may appear and expand rapidly after onset or change of TKI-therapy, and subclone-specific response to treatment can be readily documented (19). Surveillance of the size of mutant subclones during TKI-treatment was shown to provide information on their responsiveness to therapy and the presence or imminent onset of resistant disease (19–21). Although various methods for mutational analysis based on PCR-mediated pre-amplification of the BCR-ABL1 TKD permit highly sensitive subclone detection, sequence errors introduced by the reverse transcriptase and PCR polymerase must be considered (data not shown). Owing to this phenomenon, the practical detection limit of prospective screening for mutant subclones after PCR amplification of the BCR-ABL1 TKD is in the range of 1%. Although several techniques can reveal the presence of mutations even in considerably smaller cell subsets, reliable distinction between true mutations and errors introduced by the indicated enzymes is not always feasible. This restriction also applies to massively parallel sequencing using NGS, as indicated by bioinformatic analysis (22).

Recent observations suggest that occurrence of the so-called compound mutations, defined by the presence of two or more mutations on the same DNA molecule, is often associated with particularly high resistance to multiple TKIs (20, 23). The frequency of compound mutations is apparently quite high, thus rendering their reliable detection an important diagnostic challenge (20, 23). Sensitive and quantitative methods such as LD-PCR only have the capacity to identify compound mutations by the documentation of concordant kinetics of different mutant subclones in certain instances. NGS has therefore become the emerging

method of choice for sensitive sequencing of the BCR-ABL1 TKD. Due to the limited read length offered by most current NGS technologies, multiple overlapping amplicons are required to cover the entire TKD. This prevents the assignment of nucleotide substitutions located on different amplicons to the same DNA molecule, and therefore requires additional laborious steps to facilitate unequivocal identification of constellations with compound mutations (20). We have therefore addressed the possibility to overcome this disadvantage by using the GS FLX+ platform (454 Life Sciences) with an average read length of 900bp to cover the complete BCR-ABL1 TKD in a single read.

Materials and Methods

Isolation of RNA, reverse transcription, and PCR amplification of the *BCR-ABL1* TKD for LR-NGS and fragment subcloning

Primary nucleic acid analyte material was extracted from peripheral blood samples following standard total RNA extraction methods (QIAamp RNA Blood Mini Kit, Qiagen, Hilden, Germany or Eurobio, Les Ulis, France), and cDNA was synthesized using the High-Capacity cDNA Reverse Transcription Kit (Life Technologies, Carlsbad, USA) according to the manufacturers' recommendations.

For amplification of the BCR-ABL1 kinase domain, a two-step semi-nested RT-PCR was performed, as described previously (7). The Expand Long template PCR system (Roche, Basel, Switzerland) was used for amplification of PCR products for subsequent LR-NGS analysis. For the subcloning/sequencing strategy performed, an additional alternative proof-reading enzyme, the Phusion High-Fidelity DNA Polymerase (New England Biolabs, Massachusetts, USA) was applied.

Library preparation

Long-read ABL1 TKD sequencing of the BCR-ABL1 fusion gene with the GS FLX+, Roche 454 sequencing system (Life Sciences, Branford, USA) requires a specific library preparation. A semi-nested PCR (7, 10) was combined with a modified long-read specific emulsion PCR (emPCR) protocol (Technical Bulletin No. 2011-001, 454 Life Sciences) including amplicon purification, quantification and equimolar pooling of different amplicons to permit sequencing of several samples in one GS FLX+ run with comparable coverage. The pooling step requires specific labelling of the amplicons from each sample by multiplex identifiers (MIDs) to permit subsequent identification of individual sequences and their assignment to individual samples. In the first step of the semi-nested PCR (Figure 1), a reaction mix containing 0.3 μ M forward primer [5'TGACCAACTCGTGTGTGAACTC], 0.3 μ M reverse primer [5'TCCACTTCGTCTGAGATACTGGATT], 2.5 μ l 10xbuffer with 18 mM MgCl₂, 0.2 mM of each dNTPs, 1 U/ μ l Fast Start High Fidelity polymerase (Fast Start High Fidelity PCR System, Roche Applied Science, Penzberg, Germany) and 5 μ l cDNA template in a total volume of 25 μ l was run on a Verity thermal cycler (Life Technologies, Foster City, USA) with the following protocol: 5 min at 95°C, followed by 30 cycles with 30 sec at 95°C, 30 sec at 60°C, 2:30 min at 72°C, and a final extension step at 72°C for 7 min followed by cooling at 4°C.

In the second step of the semi-nested PCR (Figure 1), fusion primers were used for the production of a single amplicon covering the entire ABL1 tyrosine kinase domain of 933bp. The primers included sequences homologous to the respective ABL1 target regions of 19 bp (fwd primer) and 25 bp (rev primer), 454 adaptor sequences of 25 bp (type A and B for forward and reverse sequencing, respectively), and a MID sequence tag of 10 bp specific for each sample. One µl of the first round PCR product was added to the second PCR reaction master mix differing from the first amplification step only by the primers used (fwd: 5'CGTATCGCCTCCCTCGCGCCATCAG-MID-CGCAACAAGCCCACTGTCT; rev: 5'CTATGCGCCTTGCCAGCCCGCTCAG-MID-TCCACTTCGTCTGAGATACTGGATT). The second step of the semi-nested PCR was performed using the same thermal cycler according to the following protocol: 95°C for 5 min, 30 cycles with the profile of 20 sec at 95°C, 30 sec at 60°C, 60 sec at 72°C, and a final extension step at 72°C for 7 min followed by cooling at 4°C. PCR products were purified by Agencourt AMPure XP beads (Beckman Coulter, Vienna, Austria) using a 1:1.5 DNA bead ratio according to the provided protocol (Amplicon Library Preparation Method Manual, GS FLX Titanium Series, October 2009, 454 Life Sciences). Purified amplicons were quantified by Quant-iT PicoGreen® dsDNA Reagent (Life Technologies) utilizing an Infinite 200Pro fluorometer (Tecan, Grödig, Austria). The DNA concentration ranged from 1.11 to 1.32 ng/µl, and no differences related to the two different MIDs tested have been observed. Based on the standard concentrations, the amplicon intensity signals were directly translated from ng/µl to molecules/µl, pooled in equimolar proportions and diluted to a final concentration of 2x10E6 molecules per µl.

Long-read 454 sequencing for detection of BCR-ABL1 compound mutations

Emulsion PCR, bead recovery, bead enrichment and sequencing primer annealing were performed following the emPCR protocol (Amplification Method Manual - Lib-A MV, January 2010) according to the supplier's instructions (454 Life Sciences). For the emulsion PCR, a copy-per-bead ratio of 1 was used and carried out on a GeneAmp® PCR System 9700 (Life Technologies). Bead recovery was done manually, whereas bead enrichment and sequencing primer annealing was done with the robotic enrichment module (REMe, 454 Life Sciences) on a Microlab STARlet robot (Hamilton Robotics GmbH, Martinsried, Germany). Subsequent bead counting was performed using Scepter 2.0 (Merck Millipore, Billerica, USA), finally resulting in the loading of 790,000 beads (two samples with different MID per region) in one of four regions of the the PicoTiterPlate (454 Life Sciences). Next generation sequencing was carried out according to the appropriate sequencing method manual (Sequencing Method Manual XL+, August 2012, 454 Life Sciences) on the GS FLX+ instrument.

Data analysis

GS Sequencer Software 2.9 was used for image processing and data analysis, and took place on a high-performance computing server using the analysis pipeline Long Amplicons 2 (default, 454 Life Sciences). Generated sff-files were directly linked to the GS Amplicon Variant Analyzer (GS AVA, 454 Life Sciences) project (version 2.9) following automated demultiplexing and reference alignment (NM_005157.4 obtained from NCBI was used as a reference sequence). Alignment for variants with 1% hit percentage were examined (forward/reverse coverage, alignment problems, homopolymer adjacencies). A variant 1%

was taken as a true variant if there were no critical sequencing adjacencies, and the variant was sequenced in both directions. Sample No 8 (Table 1) showed insufficient results in the GS AVA software, and further investigation was therefore performed using Genomics Workbench 6.0.1 (CLCbio, Aarhus, Denmark). None of the two software tools was able to perform the alignment in a case displaying a 540bp deletion and a 35bp insertion. We have therefore created an artificial reference sequence excluding the 540bp deletion but including the 35bp insertion, which permitted successful alignment.

Subcloning and Sanger sequencing of the *BCR-ABL1* kinase domain

The *BCR-ABL1* TKD amplicons were ligated into pGEM®-T Easy Vector System I (Promega, Fitchburg, USA), introduced into One Shot TOP10 Chemically Competent *E. coli* (Invitrogen, Carlsbad, USA) and plated on Luria-Bertani (LB) plates containing ampicillin (100 µg/ml), IPTG (Isopropyl-β-D-thiogalactopyranoside) and X-gal (5-Brom-4-chlor-3-indoxyl-β-D-galactopyranoside). White bacterial colonies were selected and grown in LB medium supplemented with ampicillin (100 µg/ml), and plasmids were extracted using a Miniprep kit (Biozym, Hessisch Oldendorf, Germany) according to the manufacturer's recommendations.

In total, 101 clones derived from two select samples, including 47 clones from sample No 2, and 54 clones from sample No 10 (Table 1), were subjected to Sanger sequencing. The sequence data obtained were compared to the wild-type *ABL1* sequence (*Acc.Nr.* 62362413.4-3396) using the BLAST alignment tool (NCBI) and the Mutation Surveyor(demo)Version3.10 (SoftGenetics, LLC State College, USA).

Results and Discussion

Candidate patients selected for the screening of compound mutations in the *BCR-ABL1* kinase domain included individuals displaying single, multiple or no detectable point mutations by established methods such as Sanger sequencing, pyrosequencing (13) or ligation-dependent PCR (10). A total of eleven peripheral blood specimens from patients with chronic myeloid leukemia (CML) were subjected to long-range next generation sequencing using the Roche GS FLX+ instrument. The specimens investigated included individual samples obtained from four patients, and serial samples from two patients, including two and five consecutive samples, respectively (Table 1, patients 1 and 4). A protocol facilitating coverage of the entire *ABL1* kinase domain in a single read was used. The median coverage of individual target sequences was 38.401 reads, with a range from 20.388 to 59.825 reads. The reads obtained in sample No 11 (Table 1) were beyond the indicated range because sequencing was performed in one lane facilitating higher coverage (150.997 reads). Based on bioinformatic analysis of the raw data, only subclones representing 1% of the entire *BCR-ABL1* positive leukemic clone could be considered as real variants. Smaller subclones identified by the LR-NGS analysis cannot be reliably distinguished from possible artifacts introduced by the reverse transcriptase and PCR polymerase during the preparation of PCR amplicons for sequencing analysis or during the NGS process. The choice of individual enzymes may have an impact on the error rate. In the present study, we have compared two different proofreading-polymerases including the

Phusion High-Fidelity DNA Polymerase (New England Biolabs) and the Expand Long template PCR system (Roche), and the latter resulted in a considerably higher number of errors. In agreement with earlier observations, most of the erroneous nucleotide substitutions detected by subclone sequencing were transitions and only rarely transversions (data not shown). However, regardless of the enzymes used, the problem cannot be entirely eliminated, thus affecting the actual detection limit for mutations by any downstream sequencing approach. The notion that the reverse transcription step contributes the majority of occurring errors could be addressed by using DNA rather than RNA as template material. However, for *BCR-ABL1* TKD analysis, this approach would prevent the possibility of eliminating the non-rearranged *ABL1* allele from amplification, and would require exon-specific analysis. When sequencing homopolymer motifs, 454 pyrosequencing may produce false nucleotide insertions and deletions. To account for this problem, only variants indicated by 1% of the reads in both sequencing directions are called (Figure 3).

The use of NGS is a well-established approach to sensitive screening for mutant subclones in the *BCR-ABL1* kinase domain (20, 24). The limited read-length in the range of 36-400 bp currently provided by various NGS instruments is an impediment to the identification of compound mutations within larger target regions of interest. In a recently published study, four overlapping amplicons were used to cover the entire *BCR-ABL1* TKD. In the presence of mutations located on different PCR fragments, additional sequencing of specifically designed new amplicons was required to reliably identify constellations with compound mutations (20). In view of the very high current costs of NGS analysis, this approach may not be readily amenable to the screening for compound mutations in the routine setting.

In the LR-NGS approach presented, the entire *BCR-ABL1* TKD can be sequenced by a single read (Figure 1), thereby permitting easy assignment of mutations to the same or to different molecules. The observations in the patients investigated are summarized in Table 1, and indicate the presence of single and compound point mutations as well as insertions and deletions (indels) leading to frameshifts in all instances. Most of the aberrations detected were described previously, such as the single mutations p.G250E, p.Y253H, p.T315A, p.F317I, p.Q252H, p.T315I and p.L248V, the compound mutation p.G250E/p.T315A, the 35bp insertion (p.C475fs*11) at exons 8/9, the 185 bp deletion (p.R362fs*21) of exon 7 (Figure 2), and the 81 bp deletion (p.L248_K274del) at exon 4, with or without concomitant presence of other changes (25–31). Newly observed aberrations included a 540 bp deletion affecting the complete exons 7 and 8, and partially exons 6 and 9 (Table 1, sample No 8) as well as the compound mutations p.G250E/p.F317I, p.G250E/p.F317I/p.E499E and p.G250E/p.T315A/p.E499E which, to our knowledge, have not been reported previously.

To address the possibility of artifacts inherent in the technique that could lead to incorrect identification of single and compound mutations or indels, the LR-NGS findings were reevaluated by an independent technical approach in select cases (samples 2 and 10, Table 1). The PCR amplicons of the *BCR-ABL1* TKD derived from individual specimens were subcloned into pGEM®-T plasmids, and >100 clones were subjected to analysis by Sanger sequencing. The observations made by LR-NGS analysis in subclones representing 1.5% of the entire leukemic clone, including single mutations, compound mutations and combinations of point mutations with insertions or deletions (e.g. p.G250E/p.F317I,

p.G250E/p.T315A, p.G250E/p.F317I/p.E499E, p.T315I/p.R362fs*21 (Table 1), could generally be identified in individual clones by Sanger sequencing, thus documenting the reliability of the LR-NGS technology. The practical detection limit of 1% for mutational analysis in the *BCR-ABL1* TKD provided by the LR-NGS technique is identical to that reported for the more laborious and costly four-fragment NGS approach (20). The ability to sequence the entire target region of interest in a single read permits reliable and more economic assignment of mutations and other aberrations to single or different leukemic cells. The LR-NGS approach presented therefore has the potential of becoming the method of choice for mutational screening, particularly for the identification of compound mutations in Ph-positive leukemias.

Acknowledgement

This work was supported by the Austrian Science Fund (FWF), SFB grants F4705-B20 and F4704-B20.

References

1. Jabbour E, Kantarjian H. Chronic myeloid leukemia: 2012 update on diagnosis, monitoring, and management. *Am J Hematol.* 2012; 87:1037–45. [PubMed: 23090888]
2. Branford S, Melo JV, Hughes TP. Selecting optimal second-line tyrosine kinase inhibitor therapy for chronic myeloid leukemia patients after imatinib failure: does the BCR-ABL mutation status really matter? *Blood.* 2009; 114:5426–35. [PubMed: 19880502]
3. Quintas-Cardama A, Kantarjian H, Cortes J. Flying under the radar: the new wave of BCR-ABL inhibitors. *Nat Rev Drug Discov.* 2007; 6:834–48. [PubMed: 17853901]
4. Soverini S, Colarossi S, Gnani A, et al. Contribution of ABL kinase domain mutations to imatinib resistance in different subsets of Philadelphia-positive patients: by the GIMEMA Working Party on Chronic Myeloid Leukemia. *Clin Cancer Res.* 2006; 12:7374–9. [PubMed: 17189410]
5. Willis SG, Lange T, Demehri S, et al. High-sensitivity detection of BCR-ABL kinase domain mutations in imatinib-naïve patients: correlation with clonal cytogenetic evolution but not response to therapy. *Blood.* 2005; 106:2128–37. [PubMed: 15914554]
6. Gruber FX, Lamark T, Anonli A, et al. Selecting and deselecting imatinib-resistant clones: observations made by longitudinal, quantitative monitoring of mutated BCR-ABL. *Leukemia.* 2005; 19:2159–65. [PubMed: 16224487]
7. Branford S, Hughes T. Detection of BCR-ABL mutations and resistance to imatinib mesylate. *Methods Mol Med.* 2006; 125:93–106. [PubMed: 16502579]
8. Khorashad JS, Anand M, Marin D, et al. The presence of a BCR-ABL mutant allele in CML does not always explain clinical resistance to imatinib. *Leukemia.* 2006; 20:658–63. [PubMed: 16467863]
9. Vivante A, Amariglio N, Koren-Michowitz M, et al. High-throughput, sensitive and quantitative assay for the detection of BCR-ABL kinase domain mutations. *Leukemia.* 2007; 21:1318–21. [PubMed: 17330094]
10. Preuner S, Denk D, Frommlet F, Nesslerboeck M, Lion T. Quantitative monitoring of cell clones carrying point mutations in the BCR-ABL tyrosine kinase domain by ligation-dependent polymerase chain reaction (LD-PCR). *Leukemia.* 2008; 22:1956–61. [PubMed: 18432261]
11. Ernst T, Erben P, Muller MC, et al. Dynamics of BCR-ABL mutated clones prior to hematologic or cytogenetic resistance to imatinib. *Haematologica.* 2008; 93:186–92. [PubMed: 18223278]
12. Pelz-Ackermann O, Cross M, Pfeifer H, et al. Highly sensitive and quantitative detection of BCR-ABL kinase domain mutations by ligation PCR. *Leukemia.* 2008; 22:2288–91. [PubMed: 18615107]
13. Alikian M, Gerrard G, Subramanian PG, et al. BCR-ABL1 kinase domain mutations: methodology and clinical evaluation. *Am J Hematol.* 2012; 87:298–304. [PubMed: 22231203]

14. Soverini S, Hochhaus A, Nicolini FE, et al. BCR-ABL kinase domain mutation analysis in chronic myeloid leukemia patients treated with tyrosine kinase inhibitors: recommendations from an expert panel on behalf of European LeukemiaNet. *Blood*. 2011; 118:1208–15. [PubMed: 21562040]
15. Baccarani M, Cortes J, Pane F, et al. Chronic myeloid leukemia: an update of concepts and management recommendations of European LeukemiaNet. *J Clin Oncol*. 2009; 27:6041–51. [PubMed: 19884523]
16. Baccarani M, Deininger MW, Rosti G, et al. European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. *Blood*. 2013
17. Lange T, Ernst T, Gruber FX, et al. The quantitative level of T315I mutated BCR-ABL predicts for major molecular response to second-line nilotinib or dasatinib treatment in patients with chronic myeloid leukemia. *Haematologica*. 2013; 98:714–7. [PubMed: 23065514]
18. Khorashad JS, Milojkovic D, Mehta P, et al. In vivo kinetics of kinase domain mutations in CML patients treated with dasatinib after failing imatinib. *Blood*. 2008; 111:2378–81. [PubMed: 17982022]
19. Preuner S, Mitterbauer G, Mannhalter C, et al. Quantitative monitoring of BCR/ABL1 mutants for surveillance of subclone-evolution, -expansion, and -depletion in chronic myeloid leukaemia. *Eur J Cancer*. 2012; 48:233–6. [PubMed: 21955823]
20. Soverini S, De Benedittis C, Machova Polakova K, et al. Unraveling the complexity of tyrosine kinase inhibitor-resistant populations by ultra-deep sequencing of the BCR-ABL kinase domain. *Blood*. 2013; 122:1634–48. [PubMed: 23794064]
21. Gruber FX, Ernst T, Porkka K, et al. Dynamics of the emergence of dasatinib and nilotinib resistance in imatinib-resistant CML patients. *Leukemia*. 2012; 26:172–7. [PubMed: 21818112]
22. Niklas N, Proll J, Danzer M, Stabenheimer S, Hofer K, Gabriel C. Routine performance and errors of 454 HLA exon sequencing in diagnostics. *BMC Bioinformatics*. 2013; 14:176. [PubMed: 23731822]
23. Khorashad JS, Kelley TW, Szankasi P, et al. BCR-ABL1 compound mutations in tyrosine kinase inhibitor-resistant CML: frequency and clonal relationships. *Blood*. 2013; 121:489–98. [PubMed: 23223358]
24. Soverini S, De Benedittis C, Machova Polakova K, et al. Dissecting the Complexity of Philadelphia-Positive Mutated Populations by Ultra-Deep Sequencing of the Bcr-Abl Kinase Domain: Biological and Clinical Implications. *Blood (ASH Annual Meeting Abstracts)*. 2012; 120:692.
25. Cho EH, Kang SK, Kim SY. Supplemental results of the detection of splicing variant with c-ABL exon 7 deletion by direct sequencing Comment on "A recurrent splicing variant without c-ABL Exon 7 in Imatinib-resistant patients" by Curvo et al. [*Leuk. Res.*]. *Leuk Res*. 2009; 33:505–6. [PubMed: 18848355]
26. Curvo RP, Zalberg IR, Scholl V, Pires V, Moellmann-Coelho A, Moreira MA. A recurrent splicing variant without c-ABL Exon 7 in Imatinib-resistant patients. *Leuk Res*. 2008; 32:508–10. [PubMed: 17570524]
27. Gruber FX, Hjorth-Hansen H, Mikkola I, Stenke L, Johansen T. A novel Bcr-Abl splice isoform is associated with the L248V mutation in CML patients with acquired resistance to imatinib. *Leukemia*. 2006; 20:2057–60. [PubMed: 17008892]
28. Laudadio J, Deininger MW, Mauro MJ, Druker BJ, Press RD. An intron-derived insertion/truncation mutation in the BCR-ABL kinase domain in chronic myeloid leukemia patients undergoing kinase inhibitor therapy. *J Mol Diagn*. 2008; 10:177–80. [PubMed: 18276770]
29. Ma W, Kantarjian H, Yeh CH, Zhang ZJ, Cortes J, Albitar M. BCR-ABL truncation due to premature translation termination as a mechanism of resistance to kinase inhibitors. *Acta Haematol*. 2009; 121:27–31. [PubMed: 19332983]
30. O'Hare T, Zabriskie MS, Eide CA, et al. The BCR-ABL35INS insertion/truncation mutant is kinase-inactive and does not contribute to tyrosine kinase inhibitor resistance in chronic myeloid leukemia. *Blood*. 2011; 118:5250–4. [PubMed: 21908430]
31. Sherbenou DW, Hantschel O, Turaga L, et al. Characterization of BCR-ABL deletion mutants from patients with chronic myeloid leukemia. *Leukemia*. 2008; 22:1184–90. [PubMed: 18354488]

Figure 1. Semi-nested PCR amplification strategy of the entire *BCR-ABL1* kinase domain for long-range NGS analysis.

Figure 2. Representation of long-range NGS analysis of the entire *BCR-ABL1* kinase domain
 Displayed is a representative example (clone No 1; sample No 5 in Table 1) showing the results of forward (A), reverse (B) and combined forward/reverse (C) long-range NGS sequencing. Only variants identified with the Amplicon Variant Analyzer software in 1% of reads were considered. This software was implemented in the present study in addition to standard software packages for more accurate identification of false-positive findings. Mutations must be detectable in both sequencing directions in order to qualify as true results. Positions of detected aberrations are indicated in the bottom line of panel C (see also Table 1). In the long sequence reads provided by the 454 GS FLX+ system, not every sequencing reaction continues until the very end of the template molecule, but the high-quality sequencing output of the raw data did not require any additional trimming by the processing software.

The x-axis indicates the mRNA nucleotide positions of the reference sequence of the *ABL1* kinase domain (nucleotide range: 677-1498). The left y-axis indicates the size of variants in percent of reads per region (coverage); the right y-axis shows the absolute number of reads per base pair (light blue horizontal line). Individual nucleotides are indicated by colors of vertical bars (red= T, blue= C, green= A, black= G; grey block = deletion).

Figure 3. Identification of true and false positive results in 454 next generation sequencing

Three typical examples of variants are displayed. The blocks of horizontal lines represent different segments of the ABL1 kinase domain, as indicated by the nucleotide sequence positions in the bottom line. Each vertical line represents single-read sequence information derived from one amplicon molecule.

(A) True result (clone no 1; sample No 9 in Table 1): mutations are detected in both sequencing directions at 1% frequency.

(B) False positive result (clone No 1; sample No 6 in Table 1): the homopolymer problem (stretch of adenosine residues) is demonstrated by indicating a nucleotide insertion in forward sequencing (100%) which is absent in reverse sequencing reads (99.8%). The indicated mutation is therefore disqualified.

(C) False positive result (clone No 1; sample No 7 in Table 1): a nucleotide insertion following a homopolymer (stretch of cytosine residues) is indicated in forward sequencing (~78.6 %) but is not confirmed in the reverse sequencing direction (100%). The indicated mutation is also disqualified.

Table 1
Results of LR-NGS analysis

Displayed are the single and compound mutations as well as indels observed by LR-NGS analysis in 11 CML patients. The type of mutation, the amino acid (AA) change, the relative size of mutant subclones (variant %) and the absolute read count for individual variants (Hit) are indicated. Samples 1 and 2 are sequential specimens from patient No 1 (time interval: 4 months). Samples 5, 6, 7, 8, 9 are sequential specimens from patient No 4 (time intervals: 3>1.5>1.5>1 months). The remaining samples 3, 4, 10, and 11 are single specimens from patients No 2, 3, 5, and 6, respectively.

The following single mutations were detected by Sanger Sequencing (SS), LD-PCR or pyrosequencing prior to LR-NGS analysis: Sample 1: p.G250E, p.Y253H (SS); sample 2: p.G250E, p.T315A, p.F317I (SS); sample 3: p.E459K (SS); sample 4: p.M244V (SS); sample 5: p.Q252H (pyrosequencing+LD-PCR); sample 6: p.Q252H (pyrosequencing+LD-PCR); sample 7: p.Q252H (pyrosequencing+LD-PCR); sample 10: p.T315I (pyrosequencing+LD-PCR); sample 11: p.L248V, p.L248V_K274des (SS).

Pat (No)	Sample (No)	Clone (No)	Clone Variants	AA Change	Variants (%)	Hit (reads)
Pat 1	Sample 1	clone 1	c.749 G>A	p.G250E	22.0	14,956
		clone 2	c.757 T>C	p.Y253H	26.5	18,458
	Sample 2	clone 1	c.749G>A, c.943A>G, c.1497A>G	p.G250E, p.T315A, p.E499E	15.7	4,748
		clone 2	c.749G>A, c.949T>A, c.1497A>G	p.G250E, p.F317I, p.E499E	4.1	1,240
		clone 3	c.749G>A, c.1497A>G	p.G250E, p.E499E	22.7	6,864
		clone 4	c.943A>G, c.1497A>G	p.T315A, p.E499E	17.1	16,237
		clone 5	c.949T>A, c.1497A>G	p.F317I, p.E499E	3.3	3,128
		clone 6	c.1497A>G	p.E499E	45.2	26,522
Pat 2	Sample 3	clone 1	c.1375G>A, c.1423_1424ins35	p.E459K, p.C475fs*11	6.2	3,495
		clone 2	c.1375G>A	p.E459K	90.9	51,247
Pat 3	Sample 4	clone 1	c.730A>G	p.M244V	20.3	9,887
Pat 4	Sample 5	clone 1	c.756G>T, c.1086_1270del185	p.Q252H, p.R362fs*21	5.4	1,407
		clone 2	c.756G>T, c.1423_1424ins35	p.Q252H, p.C475fs*11	1.0	160
		clone 3	c.756G>T	p.Q252H	41.0	10,845
		clone 4	c.1086_1270del185	p.R362fs*21	2.7	727
		clone 5	c.888_919del32	unknown	1.6	427
	Sample 6	clone 1	c.756G>T, c.1086_1270del185	p.Q252H, p.R362fs*11	20.4	8,151
		clone 2	c.756G>T	p.Q252H	64.6	32,472
	Sample 7	clone 1	c.756G>T	p.Q252H	51.2	38,883
	Sample 8	clone 1	c.838_1378del540, c.1423_1424ins35	unknown, p.C475fs*11	98.4	49,425
	Sample 9	clone 1	c.825G>A	p.D276K	2.2	229
Pat 5	Sample 10	clone 1	c.1086_1270del185, c.1423_1424ins35	p.R362fs*21, p.C475fs*11	1.3	278
		clone 2	c.944C>T, c.1086_1270del185	p.T315I, p.R362fs*21	9.1	2,260
		clone 3	c.944C>T	p.T315I	30.6	9,778
		clone 4	c.1086_1270del185	p.R362fs*21	12.1	3,278
		clone 5	c.1423_1424ins35	p.C475fs*11	2.1	460

Pat (No)	Sample (No)	Clone (No)	Clone Variants	AA Change	Variants (%)	Hit (reads)
Pat 6	Sample 11	clone 1	c.742_822del81, c.1423_1424ins35	p.L248_K274del, p.C475fs*11	1.4	2,510
		clone 2	c.742_822del81	p.L248_K274del	92.4	133,524
		clone 3	c.742C>G	p.L248V	4.5	6,505