

**EVALUATION OF A NEXT GENERATION SEQUENCING PANEL FOR *ABL1*  
KINASE DOMAIN MUTATIONS IN PATIENTS WITH *BCR-ABL1* POSITIVE  
CHRONIC MYELOID LEUKAEMIA**



UNIVERSITY OF THE  
WITWATERSRAND,  
JOHANNESBURG

Dr. Tanya De Jager

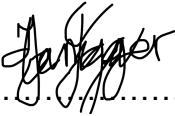
A dissertation submitted to the Faculty of Health Sciences, University of the  
Witwatersrand, in partial fulfilment of the requirements for the degree of  
MMed (Haem)

Johannesburg, 2022

## Declaration

---

I, Tanya De Jager declare that this research report is my own, unaided work. It is being submitted for the Degree of Master of Medicine at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

  
.....

(Signature of candidate)

.....17.....day of August.....20.22.....in Johannesburg.....

## **Dedication**

To my husband Graeme.

For his ever-present love and support, without which none of this would be possible.

**Presentations and publications arising from the research project**

None at present.

## Abstract

---

### Background

The commonest cause of failure to respond to imatinib or loss of molecular response in patients with *BCR-ABL1* positive chronic myeloid leukaemia (CML) is the presence of *BCR-ABL1* tyrosine kinase domain mutations. Next generation sequencing (NGS) can detect clinically relevant low-level mutations in this subgroup of patients thus guiding therapeutic options. We aimed to: 1. establish a novel “in-house” NGS assay to detect *ABL1* mutations. The Nextera™ XT DNA library preparation kit as well as the iSeq™ 100 Sequencing System were both selected for this application. The Nextera™ Kit allows for simultaneous fragmentation and tagmentation of the DNA fragment with unique sequencing adapters. This feature along with the iSeq™ 100 Sequencing System with its small size and cost effectiveness, allows single gene sequencing with less scientist hands-on time, faster TAT and reduction in all round cost for *ABL1* kinase domain sequencing. Our second aim was evaluate the performance of this assay in detecting mutations previously characterized by Sanger sequencing (SS).

### Methods

Retrospective samples from 24 patients diagnosed with CML, who underwent *BCR-ABL1* tyrosine kinase domain mutations due to suboptimal/failed/loss of response to tyrosine kinase inhibitor (TKI) therapy were tested. All samples had previously undergone Sanger sequencing. Messenger RNA (mRNA) from forty patient samples was converted to cDNA. Two rounds of polymerase chain reaction (PCR) was performed to first select-out *BCR-ABL1* and then *ABL1*. Amplification was assessed using 2% agarose gel electrophoresis for product bands that were clearly visible and of the correct size. Twenty-four samples with the best results were chosen for NGS. NGS library preparation was performed using the Illumina® (San Diego, CA, USA) Nextera™ Kit that uses transposase technology. Library quantification was performed using the Qubit™ (ThermoFisher, Waltham, USA), samples were pooled into a single tube and loaded onto the sequencing cartridge. Data analysis was initiated as the sequencing run progressed and the sequencing data aligned to the reference genome (Hg38) specifically using chromosome 9. The Illumina® (San Diego, CA, USA) variant caller software separated patient samples based on the

unique barcode, after which the software package Burrows-Wheeler-Aligner (BWA) was used for mapping low-divergent sequences against the reference genome. A binary Alignment Map (BAM) file is generated first and used to create the Variant Call File (VCF) that contains the data of the mutations detected during the sequencing run

## Results and Discussion

Due to non-specific amplification suspected to have occurred during the first round PCR, off target gene sequences, in addition to *BCR-ABL1*, were also amplified and sequenced during the first NGS run. During analysis of the samples it was noted that amplicons were aligning to numerous other sequences in the reference genome. After trouble shooting and a new set of PCR primers, the second sequencing attempt was successful. Here we describe the design of the assay, improvements to the original design and preliminary NGS results compared to Sanger sequencing results. Sequencing of *ABL1* during the second NGS run yielded good sequencing reads on 23 out of 24 samples and detected mutations that were also identified using SS. In addition, the average depth of coverage was noted to be ~5000.

## Conclusion

The Illumina® (San Diego, CA, USA) iSeq™ 100 Sequencing System for *ABL1* KD mutational analysis, using the semi nested *BCR-ABL1* PCR product was successful in detecting 23/24 mutations previously detected by SS. This study does not allow for a full validation owing to the retrospective nature of samples and a prospective assessment is planned to validate and implement NGS for this purpose.

Multiple international studies have shown NGS to be a reliable, accurate and sensitive diagnostic tool for this purpose. With both the European LeukaemiaNet and the National Comprehensive Cancer Networks recommending NGS in CML for *ABL1* KD mutational analysis, it is imperative that the SCGU laboratory, a large centre which services multiple clinical haematology oncology centres across South Africa be able to offer NGS for *ABL1* KD mutational analysis. Validation and implementation of this test will increase turnaround time, decrease scientist hands-on time, and reduce the cost per test. In addition, once suboptimal response or failure is detected, earlier detection of sensitive clones allows clinicians to switch to more appropriate therapy.



## **Acknowledgements**

---

Dr Pascale Willem – Head of Somatic Cell Genetics

Mr Mishalan Moodley – Senior Medical Scientist in the Department of Somatic Cell Genetics.

Dr Nikki Bouwer – Senior Haematologist, Department of Molecular Medicine and Haematology.

Ms Marija Kvas – Senior Applications Specialist, Separations and the Separations team for providing the kits we used for our study.

# Table of Contents

|                                                                                                                                                        |           |
|--------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| Declaration .....                                                                                                                                      | ii        |
| Abstract .....                                                                                                                                         | v         |
| Acknowledgements.....                                                                                                                                  | viii      |
| List of Figures .....                                                                                                                                  | xi        |
| List of Tables .....                                                                                                                                   | xii       |
| Nomenclature .....                                                                                                                                     | xiii      |
| <b>1 BACKGROUND AND LITERATURE REVIEW .....</b>                                                                                                        | <b>1</b>  |
| 1.1 Literature Review .....                                                                                                                            | 1         |
| 1.1.1 Chronic myeloid leukaemia epidemiology .....                                                                                                     | 1         |
| 1.1.2 Pathobiology of chronic myeloid leukaemia .....                                                                                                  | 1         |
| 1.1.3 Chronic myeloid leukaemia, the clinical setting .....                                                                                            | 2         |
| 1.1.4 Monitoring of patients with chronic myeloid leukaemia on tyrosine kinase inhibitors .....                                                        | 4         |
| 1.1.5 Resistance to tyrosine kinase inhibitors .....                                                                                                   | 7         |
| 1.1.6 When to perform <i>BCR-ABL1</i> kinase domain mutational analysis .....                                                                          | 10        |
| 1.1.7 How to perform <i>BCR-ABL1</i> kinase domain mutational analysis .....                                                                           | 11        |
| 1.1.8 Overview of next generation sequencing methods: .....                                                                                            | 15        |
| 1.2 Study aims and objectives .....                                                                                                                    | 19        |
| <b>2 METHODOLOGY .....</b>                                                                                                                             | <b>20</b> |
| 2.1 Research Question.....                                                                                                                             | 20        |
| 2.2 Materials and methods: .....                                                                                                                       | 21        |
| 2.2.1 Study Population: .....                                                                                                                          | 21        |
| 2.2.2 Sample selection and eligibility criteria:.....                                                                                                  | 21        |
| 2.2.3 RNA extraction and cDNA synthesis:.....                                                                                                          | 21        |
| 2.2.4 PCR amplification for <i>BCR-ABL1</i> :.....                                                                                                     | 22        |
| 2.2.5 Qubit assay: .....                                                                                                                               | 24        |
| 2.2.6 DNA library preparation using Illumina Nextera™ XT DNA library Preparation Kit (Illumina, CA, USA): .....                                        | 25        |
| 2.2.7 Sequencing on the iSeq™ 100 system: .....                                                                                                        | 27        |
| 2.2.8 Analysis:.....                                                                                                                                   | 28        |
| 2.3 Data analysis:.....                                                                                                                                | 29        |
| 2.3.1 Quality control of sequencing data: .....                                                                                                        | 29        |
| 2.3.2 Mapping and variant calling: .....                                                                                                               | 29        |
| 2.3.3 Second attempt at <i>BCR-ABL1</i> amplification with published primers:.....                                                                     | 31        |
| <b>3 RESULTS: .....</b>                                                                                                                                | <b>31</b> |
| 3.1 First attempt using the first PCR amplification product to sequence <i>ABL1</i> : .....                                                            | 31        |
| 3.1.1 Gel electrophoresis of initial first and second round PCR amplification: .....                                                                   | 31        |
| 3.1.2 Results of patient samples after quantification and clean-up:.....                                                                               | 32        |
| 3.1.3 Sequencing results and analysis: .....                                                                                                           | 34        |
| 3.2 Resequencing of <i>ABL1</i> using a different set of published primers and the product of the semi nested <i>BCR-ABL1</i> product as target: ..... | 38        |
| 3.2.1 Results of second sequencing run: .....                                                                                                          | 40        |
| <b>4 DISCUSSION: .....</b>                                                                                                                             | <b>42</b> |

|                                                                                                  |           |
|--------------------------------------------------------------------------------------------------|-----------|
| 4.1 Selection of iSeq™ 100 Sequencing System: .....                                              | 43        |
| 4.2 Evaluation of PCR protocols and sequencing of a retrospective cohort: .....                  | 43        |
| 4.3 Results of NGS and comparison to SS: .....                                                   | 44        |
| 4.4 Further investigations needed to the implement NGS in the routine diagnostic laboratory..... | 45        |
| 4.5 Limitations:.....                                                                            | 46        |
| 4.5.1 Potential for introduction of error: .....                                                 | 46        |
| 4.5.2 Small sample size:.....                                                                    | 47        |
| <b>5 CONCLUSION:.....</b>                                                                        | <b>47</b> |
| <b>6 REFERENCES .....</b>                                                                        | <b>48</b> |
| <b>7 Appendices: .....</b>                                                                       | <b>52</b> |
| 7.1 Appendix A: Reagents used for PCR reactions.....                                             | 52        |
| 7.2 Appendix B: Method for making an agarose gel.....                                            | 54        |
| 7.3 Appendix C: Materials used for Illumina® library preparation and sequencing. ....            | 55        |
| 7.4 Appendix D: Manufacturers' protocol (62).....                                                | 57        |
| 7.5 Appendix E: Snapshot of sequences obtained from the failed sequencing run. ....              | 61        |
| 7.6 Appendix F: Wits ethics clearance.....                                                       | 62        |
| 7.7 Appendix G: Plagiarism/turn-it-in report cover page.....                                     | 63        |
| .....                                                                                            | 63        |

## List of Figures

---

|                                                                                                                                                                                                |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1.1 Point mutations in <i>BCR-ABL1</i> KD which are poorly sensitive to Imatinib, Nilotinib, Bosutinib and Ponatinib (Adapted from Soverini et al, 2019)(43). .....                     | 10 |
| Figure 1.2 Figure illustrating the sensitivity of different molecular testing modalities for detecting <i>ABL1</i> kinase domain mutations (Adapted from Izzo <i>et al</i> , 2019) (51). ..... | 13 |
| Figure 1.3 Overview of the major steps of NGS.....                                                                                                                                             | 15 |
| Figure 1.4 Illustration of how the Nextera XT DNA library kit fragments and tags input DNA with unique adapter sequences. ....                                                                 | 17 |
| Figure 2.1 Illustration of primers used for semi-nested PCR .....                                                                                                                              | 24 |
| Figure 2.2 Diagram indicating the steps for library preparation and sequencing. ....                                                                                                           | 27 |
| Figure 2.3 Bioinformatic steps used for analysis.....                                                                                                                                          | 30 |
| Figure 3.1 Results of 2% Agarose gel electrophoresis. ....                                                                                                                                     | 32 |
| Figure 3.2 Average of the Quality Scores for the sequencing run. ....                                                                                                                          | 35 |
| Figure 3.3 Indicates the absence of sequencing of sequence adaptors during the sequencing run.....                                                                                             | 36 |
| Figure 3.4 Results of samples run on the LabChip® GX Touch. ....                                                                                                                               | 37 |
| Figure 3.5 Results of second round PCR using Soverini primers. ....                                                                                                                            | 38 |
| Figure 7.1 Snapshot of sequences obtained of chromosome 9.....                                                                                                                                 | 61 |

## List of Tables

---

|                                                                                                                                                                                             |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Table 1.1 Milestones to be reached when treating CML, expressed as ratio of <i>BCR-ABL1</i> against a reference gene, on the International Scale (taken from the ELN 2020 guidelines (16)). | 5  |
| Table 1.2 Cost analysis, comparing SS to NGS for <i>ABL1</i> KD mutational analysis.                                                                                                        | 12 |
| Table 2.1 Patient characteristics and mutations of the samples evaluated for sequencing.                                                                                                    | 22 |
| Table 2.2 Sample number with assigned unique sequence adapter.                                                                                                                              | 26 |
| Table 3.1 cDNA concentration, as measured on Qubit™ fluorometer, (ThermoFisher, Waltham, USA) before and after clean-up.                                                                    | 32 |
| Table 3.2 List of samples used for NGS processing (first sequencing run).                                                                                                                   | 34 |
| Table 3.3 Concentration of second round product for sequencing.                                                                                                                             | 39 |
| Table 3.4 Concordance between NGS and variants previously confirmed on Sanger sequencing.                                                                                                   | 40 |
| Table 7.1 PCR reagents used to generate cDNA, first round <i>BCR-ABL1</i> and second round <i>ABL1</i> .                                                                                    | 52 |

## Nomenclature

---

|                  |                                                        |
|------------------|--------------------------------------------------------|
| <i>ABCB1</i>     | ATP binding cassette subfamily B member                |
| <i>ABCG1</i>     | ATP binding cassette transporter 1                     |
| <i>ABL1</i>      | Abelson Murine Leukaemia 1                             |
| ACA              | Additional Cytogenetic Abnormalities                   |
| ALL              | Acute Lymphoblastic Leukaemia                          |
| AP               | Accelerated Phase                                      |
| <i>ASXL1</i>     | Additional Sex Combs-like 1                            |
| ATP              | Adenosine Triphosphate                                 |
| BAM file         | Binary Alignment Map file                              |
| <i>BCR</i>       | Breakpoint Cluster Region                              |
| BP               | Blast Phase                                            |
| BWA              | Burrows-Wheeler-Aligner                                |
| CBA              | Chromosome G Banding Analysis                          |
| cDNA             | Complementary Deoxyribonucleic acid                    |
| CF               | Conversion Factor                                      |
| CML              | Chronic myeloid leukaemia                              |
| CP               | Chronic Phase                                          |
| dNTP             | Deoxyribonucleotide triphosphate                       |
| ELN              | European Leukaemia Nt                                  |
| ELTS             | EUTOS Long-Term Survival                               |
| EUTOS            | European Treatment and Outcome Study                   |
| FISH             | Fluorescence in-situ hybridisation                     |
| HOU              | Head of Unit                                           |
| <i>IDH1/IDH2</i> | Isocitrate dehydrogenase 1/ Isocitrate dehydrogenase 2 |
| IS               | International Standard                                 |
| ISP              | Ion Sphere™ Particle                                   |
| <i>Jak2</i>      | Janus kinase 2                                         |
| KD               | Kinase Domain                                          |

|             |                                                  |
|-------------|--------------------------------------------------|
| MCyR        | Major Cytogenetic Response                       |
| MMR         | Major Molecular Response                         |
| NCCN        | National Comprehensive Cancer Network            |
| NGS         | Next Generation Sequencing                       |
| NHLS        | National Health Laboratory Service               |
| NTC         | Non-Template Control                             |
| MR          | Molecular Remission                              |
| MMR         | Major molecular response                         |
| OCT1        | Organic cation transporter 1                     |
| PCR         | Polymerase Chain Reaction                        |
| P-gp        | P-glycoprotein                                   |
| Ph+         | Philadelphia positive                            |
| RNA         | Ribonucleic acid                                 |
| RQ-PCR      | Real-Time Quantitative Polymerase Chain Reaction |
| RT-QPCR     | Reverse Transcriptase Polymerase Chain Reaction  |
| SA          | South Africa                                     |
| SAM file    | Sequence alignment map file                      |
| SLC22A1     | Solute Carrier Family 22 Member 1                |
| SNP         | Single nucleotide polymorphism                   |
| SCGU        | Somatic Cell Genetics Units                      |
| SS          | Sanger Sequencing                                |
| TAT         | Turn Around Time                                 |
| TBE         | Tris-Borate-Ethylenediaminetetraacetic acid      |
| <i>TET2</i> | tet methyl cytosine dioxygenase 2                |
| TFR         | Treatment Free Remission                         |
| TKI         | Tyrosine Kinase Inhibitor                        |
| VAF         | Variant allele frequency                         |
| VCF file    | Variant Call Format file                         |
| WHO         | World Health Organisation                        |

## 1 BACKGROUND AND LITERATURE REVIEW

---

### 1.1 Literature Review

#### 1.1.1 Chronic myeloid leukaemia epidemiology

Chronic myeloid leukaemia (CML) is a clonal myeloproliferative neoplasm characterized by a predominance of granulocytes in bone marrow and peripheral blood. According to the revised World Health Organisation (WHO) classification of tumours of haematopoietic and lymphoid tissues, the annual world-wide incidence of CML is 1-2 cases per 100 000 population. The incidence can be further divided into the occurrence among children, <0.1 cases per 100 000, to the occurrence in adults which is approximately  $\geq 2.5$  per 100 000 in elderly individuals. A slight male predominance has also been described (1). CML statistics for South Africa are lacking as the South African Cancer Registry does not distinguish between different types of leukaemia. Leukaemia data from the last available registry of 2019, shows a crude incidence of 0.92/100 000 for females and 1.04/100 000 for males per year. Furthermore, the number of new cases reported for known gender were 275 new cases for females and 299 leukaemia cases for males. These statistics however do not help us discern how many of these cases are CML (2).

#### 1.1.2 Pathobiology of chronic myeloid leukaemia

The discovery of the Philadelphia (Ph) chromosome, a shortened chromosome 22, subsequently shown (1973) to derive from the translocation  $t(9;22)(q34;q11)$ , and its consistent presence in patients with CML was one of the first chromosomal abnormalities associated with malignancy when present (3, 4). The translocation of chromosome 9 and chromosome 22, leads to the fusion of two individual genes, Abelson murine leukaemia 1 gene (*ABL1*) located on chromosome 9 with the breakpoint cluster region gene (*BCR*) on chromosome 22. Individually, each of these genes possess no intrinsic oncogenic properties, however once combined by the translocation, a fusion protein, *BCR-ABL1* is formed (5). This chimeric protein has constitutive tyrosine kinase activity, which triggers multiple downstream signalling pathways and promotes the development of chronic phase CML. This happens

through uninhibited proliferation of transformed cells, discordant maturation, escape from apoptosis, an altered bone marrow microenvironment. It further contributes to leukemogenesis by causing cells to become independent of growth factors and cytokines (6).

Different *BCR-ABL1* fusion proteins result from the Ph chromosome, dependent on the breakpoint location within the *BCR* gene. Three distinct protein products are noted. Most patients diagnosed with CML express the p210<sup>*BCR-ABL1*</sup>. The breakpoint occurs between exons 12-16 of *BCR*, a region termed the Major breakpoint cluster region. A third of patients with Ph-positive B cell acute lymphoblastic leukaemia (ALL), also have this breakpoint. The second most common protein product is the p190<sup>*BCR-ABL1*</sup> with a breakpoint occurring between exons 1-2 of *BCR* in the Minor breakpoint cluster region. It occurs in an estimated two thirds of Ph-positive B-cell ALL and an extremely small percentage of CML patients. The rarest chimeric protein product is called the p230<sup>*BCR-ABL1*</sup> with a breakpoint between exons 19-20 of *BCR*, also known as the Micro breakpoint cluster region, often seen in patients with chronic neutrophilic leukaemia (7, 8).

### **1.1.3 Chronic myeloid leukaemia, the clinical setting**

#### **1.1.3.1 Clinical presentation**

Most patients diagnosed with CML are asymptomatic and diagnosed as part of a routine blood investigation for unrelated medical complaints and then found to have a high white cell count. Ninety five percent of patients diagnosed with CML, are in chronic phase (CP). Common symptoms are unintentional weight loss, fatigue, night sweats and splenomegaly. Symptoms of splenomegaly are anorexia or early satiety due to the spleen compressing into the stomach (9).

#### **1.1.3.2 Diagnosis of chronic myeloid leukaemia.**

The natural course of untreated CML is triphasic. Patients are typically diagnosed in chronic phase. This phase is characterized by a high white cell count, which consists of predominantly granulocytes and myelocytes, usually with increased basophils and eosinophils. If left untreated, patients evolve from chronic phase to accelerated phase.

Myelopoiesis becomes ineffective with a blast count that increases in either the bone marrow or the peripheral blood and ranges from 10-19% (1, 9). The third phase (blast phase) is characterized by bone marrow failure and a blast count of more than 20%. Once a patient has reached blast phase, the disease clinically resembles an acute leukaemia, and the prognosis is generally poor.

The diagnosis of CML is typified by the detection of the *BCR-ABL* fusion gene, detected by either cytogenetic analysis, fluorescence in-situ hybridization (FISH), or via reverse transcriptase quantitative polymerase chain reaction (RT-QPCR). The respective findings should be correlated with morphological and clinical features.

### **1.1.3.3 Treatment of chronic myeloid leukaemia.**

The development of tyrosine kinase inhibitors (TKI's), particularly imatinib, revolutionised the treatment and course of CML. Imatinib (the first tyrosine kinase inhibitor) introduced onto the market in 2002 showed huge success in treating CML, and became the standard of care in the management of CML. With implementation of Imatinib in the treatment of CML, mortality decreased from 10-20% to 1-2% (4). No targeted therapy existed for the treatment of CML until the introduction of imatinib. Chemotherapeutic agents such as busulfan, hydroxyurea, and Interferon were the mainstay of treatment. Even though these drugs were able to increase the survival of CML patients, confounding factors that hindered the use of these drugs included toxicity and efficacy in controlling the disease. Allogeneic stem cell transplant is the only curative treatment for CML, however it carries significant morbidity and mortality, expense and the need for an appropriately matched donor (10). Owing to the excellent outcomes seen with TKI's, one could argue that TKI's provide a functional cure for patients with CML.

Imatinib works via competitive inhibition of the adenosine triphosphate (ATP) binding site of the *BCR-ABL1* oncoprotein. It prevents further phosphorylation of downstream signalling molecules which drives cellular proliferation. Imatinib is also able to bind to the platelet derived growth factor receptor and to the c-Kit tyrosine kinase receptor (11).

With ongoing use of imatinib, resistance is seen in some patients due to the Philadelphia-positive cells (Ph+) gaining the potential to evade or become desensitised to the inhibiting effects of imatinib. Point mutations develop within the kinase domain of *ABL1* in the *BCR-ABL1* oncoprotein. Imatinib resistance led the way in the search and development of more diverse and specific agents such as the second-generation tyrosine kinase inhibitors, e.g., dasatinib and nilotinib (12).

In the case of dasatinib, a drug originally used as salvage therapy, development and testing showed its potency to be 350 times that of Imatinib in *in vitro* experiments (13). In clinical trials, testing dasatinib as front-line therapy, patients on dasatinib achieved complete cytogenetic response in an average period of 12 months more often than those patients assigned to the imatinib arm of the study. Furthermore, patients on this study were followed up for an additional period of 5 years on dasatinib therapy and showed deeper molecular responses at earlier time points than those on imatinib (5, 14). In addition to resistance to TKI therapy other common causes of failure to respond to therapy include poor compliance to treatment as well as intolerance to the medication.

#### **1.1.4 Monitoring of patients with chronic myeloid leukaemia on tyrosine kinase inhibitors**

Monitoring patients on TKI's is important for the assessment of adequate responses to therapy. Three major responses can be assessed: haematologic, cytogenetic, and molecular responses.

Using the 2020 European LeukaemiaNet guidelines (ELN), haematologic response (to be monitored every 2 weeks until a complete haematological response is achieved), is defined as the normalization of peripheral blood counts, and is used initially during the start of therapy. A complete haematological response is defined as a white cell count of  $<10 \times 10^9/L$ , no immature granulocytes present in the peripheral blood, basophils  $<5\%$ , a platelet count of  $<450 \times 10^9/L$ , and no palpable spleen on clinical examination (15).

Cytogenetic response may be useful however according to the 2020 ELN guidelines this test alone is not sufficient enough to monitor response to therapy (16)

Molecular response is defined by a reduction of *BCR-ABL1* fusion transcripts, as measured by real-time quantitative reverse transcription polymerase chain reaction (RQ-PCR). Molecular remission defined in terms of the size of the molecular clone, is assigned a value on the International Scale (IS). According to the depth of response, patients are classified according to molecular remission (MR) based on the log reduction that is measured from a standardized baseline (17, 18). Based on the amount of detectable disease, a MR result is then assigned to each patient. Major molecular response (MMR) is defined as a 3-log reduction in *BCR-ABL1* amounting to  $\leq 0.1\%$  *BCR-ABL1* on the International Scale (19). MR4 is achieved if  $\leq 0.01\%$  *BCR-ABL1*IS and MR4.5 if  $\leq 0.0032\%$  *BCR-ABL1*IS (18). The IS was developed to allow standardisation of measuring *BCR-ABL1* transcripts amongst different laboratories around the world. In short, sample exchange between a testing laboratory and a reference laboratory allows for the calculation of a conversion factor (CF). Thus, testing laboratories can express RQ-PCR results on the IS, without having to change the method used in order to obtain the result. This allows comparison of the same patient result, performed at different laboratories (20-22).

Measuring the quantity of *BCR-ABL1* transcripts as a ratio against a control gene of choice is the current gold standard for ongoing monitor of response to therapy. Both the ELN and the National Comprehensive Cancer Network (NCCN) provide milestones that must be reached in order to define whether patients are responding adequately to therapy (16, 23).

**Table 1.1 Milestones to be reached when treating CML, expressed as ratio of *BCR-ABL1* against a reference gene, on the International Scale (taken from the ELN 2020 guidelines (16)).**

|                 | <b>Optimal</b> | <b>Warning</b>                            | <b>Failure</b>                             |
|-----------------|----------------|-------------------------------------------|--------------------------------------------|
| <b>Baseline</b> | NA             | High-risk ACA,<br>high-risk ELTS<br>score | NA                                         |
| <b>3 months</b> | $\leq 10\%$    | $> 10\%$                                  | $> 10\%$ if confirmed<br>within 1-3 months |
| <b>6 months</b> | $\leq 1\%$     | $> 1-10\%$                                | $> 10\%$                                   |

|                  |              |                                                      |                                              |
|------------------|--------------|------------------------------------------------------|----------------------------------------------|
| <b>12 months</b> | $\leq 0.1\%$ | $>0.1-1\%$                                           | $>1\%$                                       |
| <b>Any time</b>  | $\leq 0.1\%$ | $>0.1-1\%$ , loss of $\leq 0.1\%$ (MMR) <sup>a</sup> | $>1\%$ , resistance mutations, high-risk ACA |

For patients aiming at Treatment Free Remission (TFR), the optimal response (at any time) is *BCR-ABL1*  $\leq 0.01\%$  (MR<sup>4</sup>).

A change of treatment may be considered if MMR is not reached by 36-48 months.

NA not applicable, ACA additional chromosome abnormalities is Ph+ cells, ELTS EUTOS long term survival score.

<sup>a</sup>Loss of MMR (*BCR-ABL1*  $>0.1\%$ ) indicates failure after TFR.

With the advent of treatment free remission (TFR) in highly selected CML patients, it is necessary to move towards clinically relevant sensitive methods of detection of measurable residual disease. TFR can be defined as discontinuation of TKI therapy in patients with CML in CP who achieve sustained deep molecular response with no evidence of relapse (24). Strict requirements are defined by both the LeukaemiaNet and NCCN that need to be met in order for a patient to qualify for TFR. These requirements state that for a patient to be considered for TFR they must 1) have chronic phase disease, 2) have been on TKI therapy for at least 3 years and, 3) have had a continued deep molecular response (defined as a *BCR-ABL1* level of  $<0.01\%$ ) lasting at least 2 years (16, 23).

TKI therapy allows patients with CML to have an almost normal life expectancy however patients on TKI's continuously experience a myriad of TKI side-effects. These include diarrhoea, muscle cramps, fluid retention, cardiovascular and pulmonary side-effect. One third of patients on TKI therapy experience moderate to severe side-effects and these side-effects tend to remain persistent over time (25, 26).

To allow patients to transition into TFR and to discontinue TKI therapy, the following ELN international recommendations have been advised for safe monitoring supervision of TFR. The recommendations have been further subdivided into "Mandatory", "Minimal" and "Optimal" criteria (16).

### **Mandatory:**

- CML in first CP only (data are lacking outside this setting)

- Motivated patient with structured communication
- Access to high quality quantitative PCR using the International Scale (IS) with rapid turn-around of PCR test results
- Patient's agreement to more frequent monitoring after stopping treatment.  
This means monthly for the first 6 months, every 2 months for months 6–12, and every 3 months thereafter.

**Minimal (stop allowed):**

- First-line therapy or second-line if intolerance was the only reason for changing TKI
- Typical e13a2 or e14a2 BCR–*ABL1* transcripts
- Duration of TKI therapy >5 years (>4 years for 2GTKI) • Duration of DMR (MR<sup>4</sup> or better) >2 years
- No prior treatment failure

**Optimal (stop recommended for consideration):**

- Duration of TKI therapy >5 years
- Duration of DMR > 3 years if MR<sup>4</sup>
- Duration of DMR > 2 years if MR<sup>4.5</sup>

### 1.1.5 Resistance to tyrosine kinase inhibitors

Despite the development and implementation of TKI's, not all patients respond favourably to treatment, with some patients showing suboptimal molecular response and others failing treatment protocols completely. Several mechanisms of resistance to TKI's have been identified and/or postulated, the most important of which is the development of mutations in the kinase domain of *ABL1* (27-29).

Mutations within the *ABL-1* kinase domain of *BCR-ABL1* may render Ph<sup>+</sup> cells resistant to imatinib first line therapy. Resistance can be primary, which is seen in patients at presentation, who never show response to therapy in contrast to acquired mutations which occur during the course of the disease. If the T315I mutation is present in *ABL1* of *BCR-ABL1*, these patients are resistant to both first and second line TKI's. As mentioned above, imatinib works by binding to the ATP binding site

within the *ABL-1* kinase domain and renders it unavailable to ATP molecules for binding thus preventing further downstream signalling from taking place. The presence of the T315I mutation, which is due to Threonine at position 315 being changed to Isoleucine, causes disruption of the H bond interaction that takes place when imatinib binds within the kinase domain of *ABL-1*. Not only does the isoleucine side chain cause hindrance to binding by imatinib, it also renders the binding site completely insensitive to imatinib and all second generation TKI's, dasatinib, nilotinib and bosutinib (6, 30).

Resistance mutations are not caused by TKI therapy, but clones may emerge due to selective pressure in response to sequential TKI therapy (31). With more options available for the treatment of CML, the assessment of the *BCR-ABL1* kinase domain mutations informs clinical practice and guides the choice of TKI's therapy most likely to provide a better clinical outcome in individual patient.

The clinical importance of kinase domain mutational analysis is illustrated in the case of patients positive for F317L/V/I/C, T315A, or V299L *ABL1* mutations. These patients respond better to treatment with nilotinib than dasatinib, and patients who are positive for Y253H, E255K/V, and F359V/I/C mutations respond better to dasatinib versus nilotinib. As mentioned above, the presence of the T315I mutation confers resistance to imatinib, nilotinib, and Dasatinib (see Figure 1.1). These patients benefit from being treated with the third generation TKI, ponatinib (12, 32).

Other causes of resistance have been studied and are briefly described. Intracellular availability of TKI's is partly dependent on the ATP-binding cassette transporter family B1 (ABCB1). Among these transporters is the multidrug resistant gene product P-glycoprotein (P-gp), as well as the breast cancer resistance protein (ABCG2). These transporters can restrict the uptake of Imatinib into the cell by actively causing efflux of the drug out of the cell. Not only are these transporters expressed on haemopoietic cells, but they can also be found on other tissues such as the testes, brain, small bowel, and canalicular membranes of hepatocytes. Their presence at above mentioned sites further increase resistance as uptake in these tissues are decreased. In the setting of CML, it has been shown that overexpression of ABCB1 is associated with a lack of Major Cytogenetic Response (MCyR). Furthermore, patients that had

overexpression of ABCB1 also showed an increased progression to advanced phase disease (33, 34). An additional transporter that has been identified which contributes to TKI resistance, is the human organic cation transporter (OCT-1). Polymorphisms in the *SLC22A1* gene, responsible for encoding the OCT-1 protein, have been identified. These polymorphisms lead to OCT-1 preventing the influx of Imatinib into the cells (35-37).

Additional cytogenetic abnormalities are defined as chromosomal abnormalities arising in the Ph<sup>+</sup> population after presentation. Some of these genetic abnormalities are indicative of disease progression and are present in approximately 80% of patients showing advanced disease. The most common genetic abnormalities are an additional Ph<sup>+</sup> chromosome, trisomy 8, and isochromosome 17q (38). The proposed mechanism of increased resistance during clonal evolution is an increase in genetic instability (39).

Further mechanisms include persistence of leukemic stem cells, which due to their quiescent state, are able to maintain disease, as well as have inherent mechanisms to cause efflux of TKI's and increased expression of *BCR-ABL1* during advanced disease (40).

Several somatic mutations in other genes have recently been identified within the Ph positive clone and are prevalent in the setting of accelerated phase CML, as well as blast phase CML. The exact contribution gene mutations are still to be defined, however most of them contribute to epigenetic modification affecting proteins that regulate DNA methylation and others that regulate post-translational histone modification. Some of these genes are tet methyl cytosine dioxygenase 2 (*TET2*), isocitrate dehydrogenase 1 and 2 (*IDH1/IDH2*), additional sex combs-like 1 (*ASXL1*), and Janus kinase 2 (*Jak2*) (41).

Several small studies looked at patients who did not reach molecular remission or transformed to accelerated or blast phase CML. These studies found that the above mutations are often present in these cases and may occur either singly or with multiple of these genes mutated (42). Further research is ongoing to determine the prognostic and therapeutic impact that these mutations have in the setting of CML (42).

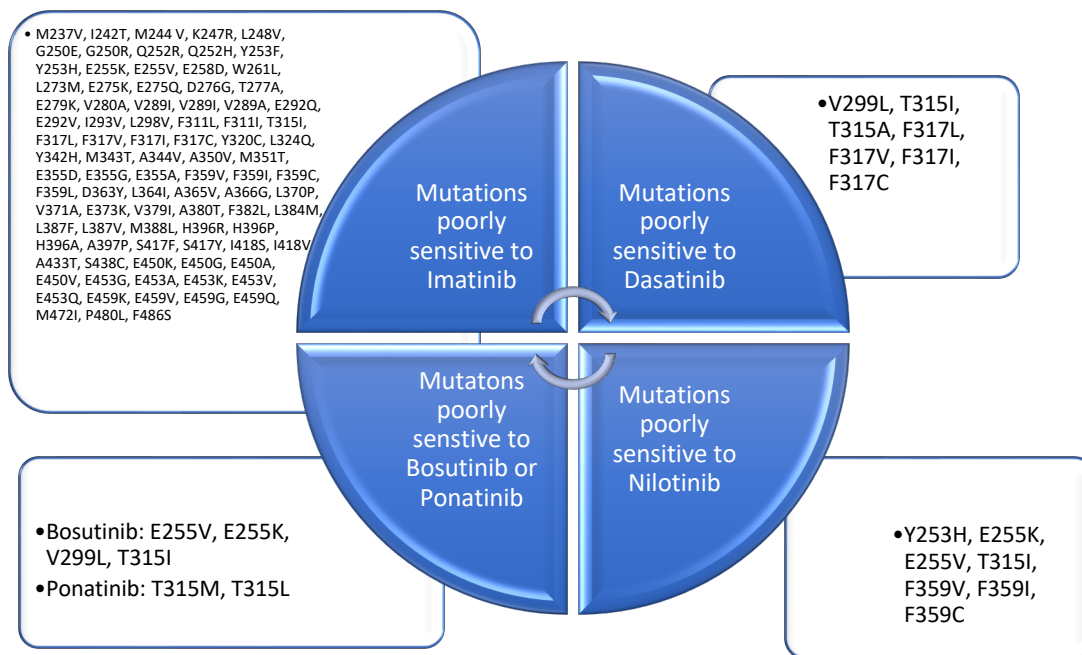


**Figure 1.1 Point mutations in *BCR-ABL1* KD which are poorly sensitive to Imatinib, Nilotinib, Bosutinib and Ponatinib (Adapted from Soverini et al, 2019)(43).**

### 1.1.6 When to perform *BCR-ABL1* kinase domain mutational analysis

Guidelines from the ELN (refer to Table 1.1) do not advise routine assessment of *BCR-ABL1* kinase domain mutations in newly diagnosed CML patients, except on the rare occasion when patients present in accelerated phase (AP) or blast phase (BP) (12). Studies have not shown mutations to be detectable at diagnosis in CP patients, even with the use highly sensitive methods, due to the low number of leukaemia-initiating cells. In the rare cases where mutations may be present, they are most likely below the detection limit of Next Generation Sequencing (NGS) (43, 44). Routine monitoring of the *BCR-ABL1* kinase domain in patients on imatinib is also not advised. The International Randomized study of Interferon and STI571 showed that 84% of patients on imatinib achieved complete cytogenetic response and never had loss of response on treatment. These patients therefore did not warrant routine assessment of *BCR-ABL1* kinase domain mutations (45).

Once a patient has been on imatinib for 12 months and shows suboptimal response (see Table 1.1) or failure to achieve molecular response, *BCR-ABL1* kinase domain analysis is indicated. Overall, 29% of patients in chronic phase CML, who fail first line imatinib therapy, show various point mutations in the *BCR-ABL1* kinase domain and are thought to be the main reason for treatment failure in these patients (12). Detection

of relevant mutations assists clinicians in deciding whether a patient's therapy needs to be changed to second generation TKI's or whether the dose of imatinib simply needs to be increased (refer to Table 1.1 for definitions of warning and failed responses).

The ELN have well established guidelines as to when to perform sequencing for CML patients, these guidelines include the following scenarios (43):

- If a patient on first line Imatinib loses major molecular response (*BCR-ABL1* >0.1%).
- Patients in chronic phase CML after first line TKI therapy with a warning response (*BCR-ABL1*IS: >10% at 3 months; >1-10% at 6 months; >0.1-1% at 12 months or any time thereafter). Testing is to be carried out regardless of the generation of TKI used. NGS may be able to detect emerging clones much earlier than SS and allow for timeous switch to a new drug.
- In patients with chronic phase CML who failed first-line therapy, regardless of the generation of TKI used. NGS can detect complex patterns of mutations including low-level mutations as well as compound mutations. This may allow individualized therapy to be instituted.
- In patients on second-line therapy where a warning or failure response is present. NGS may detect mutations that are not detectable by Sanger Sequencing (SS), which would allow for the most appropriate drug switch to be instituted.

Currently no recommendations are available for patients who are on third-line or later TKI therapy. This is due to lack of adequate definitions of optimal, warning or failure responses in these settings (43).

### **1.1.7 How to perform *BCR-ABL 1* kinase domain mutational analysis**

Up until recently, the gold standard and recommended method for analysis of *BCR-ABL1* kinase domain mutational analysis was conventional SS. SS for *ABL1* kinase domain mutations has limitations, the method is labour intensive and has limited sensitivity, clonal mutations that occur below a detection level of 20% cannot be

identified (46). The 2020 ELN guidelines and the NCCN guidelines now recommend using NGS for detection of *ABL1* kinase domain mutations (16, 23).

The cost of NGS has been steadily decreasing in recent times, and thus, is being utilised more often in a diagnostic setting. Several recent NGS studies in CML patients have shown reliable detection of emerging *BCR-ABL1* mutations, previously undetected with SS (46-48) (See Table 1.2 below for a full breakdown of cost analysis of SS vs NGS).

**Table 1.2 Cost analysis, comparing SS to NGS for *ABL1* KD mutational analysis.**

| <b>Cat no:</b>                                         | <b>Product name:</b>                                                                       | <b>Price:</b>          | <b>Price per sample:</b>            |
|--------------------------------------------------------|--------------------------------------------------------------------------------------------|------------------------|-------------------------------------|
| <b>P01B1067</b>                                        | Paxgene tube                                                                               | R11 000 (100 tubes)    | <b>R110</b>                         |
| <b>P01B9722</b>                                        | RNA extraction (Chemagic 360)                                                              | R36 000 (240 samples)  | <b>R150</b>                         |
| <b>P01B9578</b>                                        | cDNA Synthesis (Superscript VI)                                                            | R 7047,18              | <b>R140,94</b>                      |
|                                                        | <i>BCR-ABL1</i> PCR primers with ILMN tag                                                  | R 552 (8000nm)         | <b>R0.14 (2 primers)</b>            |
| <b>20018704</b>                                        | Illumina DNA Prep (24 samples)                                                             | R22 381,80             | <b>R932,58</b>                      |
| <b>20018707</b>                                        | Nextera™ DNA CD Indexes (24 Indexes, 24 Samples)                                           | R2 673,50              | <b>R111,40</b>                      |
| <b>20031371</b>                                        | iSeq 100 i1 Reagent v2 (300-cycle)                                                         | R12 061,30             | <b>R1 507,66</b>                    |
| <b>15017872</b>                                        | PhiX Control v3                                                                            | R4 190,80              | <b>R27,94</b>                       |
| <b>PO2F0459</b>                                        | 10-100ul pipette tips                                                                      | R873.74 (960 tips)     | <b>R 14.56 (R0.91 per tip x 16)</b> |
| <b>P02G0088</b>                                        | Gloves                                                                                     | R50 (100)              | <b>R3 (6 pairs per run)</b>         |
| <b>PO1B9574</b>                                        | Qubit consumables High sensitivity                                                         | R3 040,20 (500 units)  | <b>R6,08</b>                        |
| <b>CLS138948</b>                                       | LabChip consumables                                                                        | R 20 460 (750 samples) | <b>R27,28</b>                       |
| <b>Total cost per sample using iSEQ 100</b>            | (all expenses listed above)                                                                |                        | <b>R3 031,58</b>                    |
| <b>Current cost per sample using Sanger sequencing</b> | Tariff code:<br>1971 x1 (DNA extraction) – R648.10<br>1963 x4 (Genotype per PCR) – R863.80 |                        | <b>R4 103.30</b>                    |

In addition to cost saving early detection of clinically relevant clones allow clinicians to change therapy earlier in the course of disease and select the most appropriate therapy. With regards to CML and *BCR-ABL1* kinase domain mutational analysis several research groups have set up an NGS assay, specifically directed at *BCR-ABL1* kinase domain mutational analysis (46, 48, 49).

Of interest, *BCR-ABL1* mutational analysis using NGS in patients with resistance to multi-TKI therapy and with either CML or Ph+ ALL, revealed a more complex mutational status of the *BCR-ABL1* KD than initially anticipated (46). In addition, NGS can detect any sequence variation down to 3% (See Figure 1.2) thus significantly increasing the sensitivity of detection compared to SS (20%) (50).

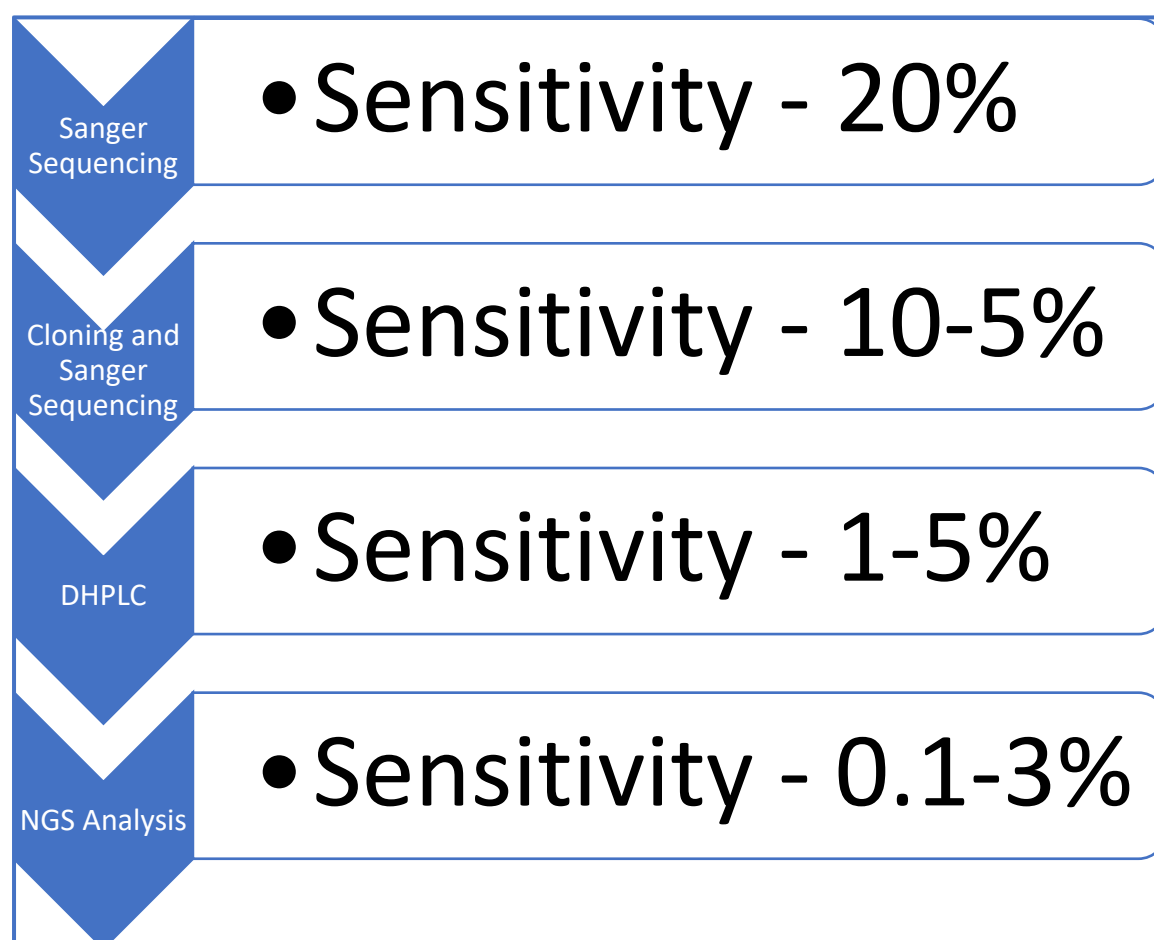


**Figure 1.2** Figure illustrating the sensitivity of different molecular testing modalities for detecting *ABL1* kinase domain mutations (Adapted from Izzo *et al*, 2019) (51).  
DHPLC: Denaturing High Performance Liquid Chromatography

In a retrospective study design from Soverini *et al*, 60 samples from patients known with CML who were found to be Imatinib resistant were selected. Two objectives were assessed: The first was to assess if NGS could track resistant mutations back to the sample of Imatinib resistance in patients who subsequently failed second generation TKI's. The second aim was to assess if low-level mutations detected at the time of Imatinib resistance, underwent clonal evolution during the course of second generation TKI therapy. As a control, samples from 24 known CML patients who were Imatinib resistant but showed good molecular response to second generation TKI's were sequenced (46).

Data showed 43% of patients who failed second generation TKI's after switch from Imatinib, already had mutations associated with failure to second line therapy at the time of switchover. These mutations were not detected during initial assessment by SS. Furthermore, all mutations with low level representation, detected at the time of switchover and known to be insensitive to second generation TKI, underwent clonal expansion. The results from this study showed that NGS is more sensitive in detecting low level clones present at the time of initial Imatinib failure as well as detecting polyclonal and compound mutations (46).

A recent paper by Soverini *et al* (43) recommends that NGS be used for sequencing in CML patients for the following reasons (refer to Table 1.1 for response definitions):

- For patients in CP CML, on first line TKI therapy, regardless of the generation of TKI used, with a “warning response”. In this scenario NGS may be able to detect low level emerging clones that are resistant to TKI, earlier than that detectable by SS. This allows for a more timeous switch to an appropriate TKI.
- In patients in CP CML, on first line TKI therapy with a “failure response”. NGS may be able to detect compound mutations which may also include low level and complex mutations. Early detection allows for more individualised therapeutic strategies.
- In patients on second line TKI therapy that show “warning or failure” response, NGS may be able to detect mutations that are not detectable via SS and allow for the appropriate change in therapy (43).

Other studies using NGS have discovered the following:

- Emerging resistant mutations were detected 2–11 months earlier than conventional SS (47) thus allowing clinicians to switch therapy early and increase survival in these patients.
- NGS is able to detect TKI-resistant mutations during major or deep molecular response (47). These mutations cannot currently be detected by SS.
- In addition to *ABL1* mutations, other mutated genes were also detected at diagnosis. These include *ASXL1*, *TET2*, *Jak2*, *IDH1* and *IDH2* (42), the clinical significance of which still needs to be defined.

#### 1.1.8 Overview of next generation sequencing methods:

A very broad overview of the major steps involved in NGS is illustrated below in Figure 1.3.

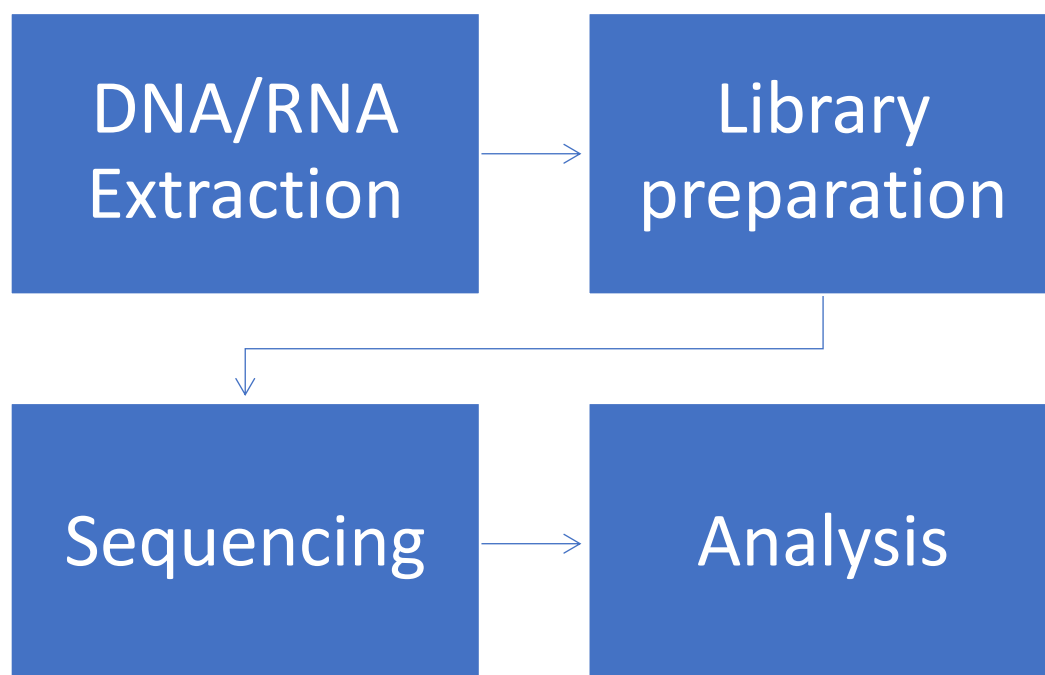


Figure 1.3 Overview of the major steps of NGS.

#### **1.1.8.1 Library preparation:**

NGS has revolutionised genomic investigations as it allows for analysis of multiple sequences in a single assay at high level of depth (depth of coverage refers to the number of times a nucleotide is read during a sequencing run). Different methodologies for sequencing are available on the market however the main principles remain the same regardless of the sequencing technology employed. In South-Africa the two main technologies used for short fragment sequencing are Illumina® (San Diego, CA, USA) and ThermoFisher® (Waltham, USA).

The principles of NGS rely on library enrichment of DNA targets either by clonal amplification of PCR produced amplicons or by hybridisation capture to generate DNA templates at high densities (52). Each targeted fragment is then sequenced in a step wise manner consisting of adding a single nucleotide in each sequencing cycle. Nucleotides complementary to the denatured target will be thus incorporated. Depending on the platform used, addition of each specific nucleotide is then detected by the analyser using either pH monitoring (Ion Torrent™, ThermoFisher (Waltham, USA)), or fluorescence detection of fluorescently labelled nucleotides (Illumina® (San Diego, CA, USA)).

The first step in sample preparation is key to a good quality sequencing run. To sequence a sample, the starting material has to be double-stranded DNA. RNA can be used, however it is first converted to cDNA before generating a library (52).

A NGS library is a collection of similarly sized DNA fragments with known adapter sequences added to the 5' and 3' ends of the target sequences respectively. Library preparation generally has three main steps. As mentioned above, the first step is DNA target selection performed by either generating an amplicon-based library using PCR primers specific for the gene of interest, or target hybridisation capture.

Fragmentation of DNA fragments is performed to generate target specific short fragments, compatible in size to the respective platform used for sequencing. Fragmentation is performed enzymatically utilising a PCR step that digests unbound

fragments or by transposases that cut DNA fragments approximately every 150-200bp (refer to Figure 1.4). Our unit specifically chose the Illumina® (San Diego, CA, USA) Nextera® XT DNA library preparation kit that utilises transposase technology. This method allows for DNA to be cut whilst barcodes are added in tandem, in a single tube reaction. This method enables a fast workflow and library preparation within a few hours (53).

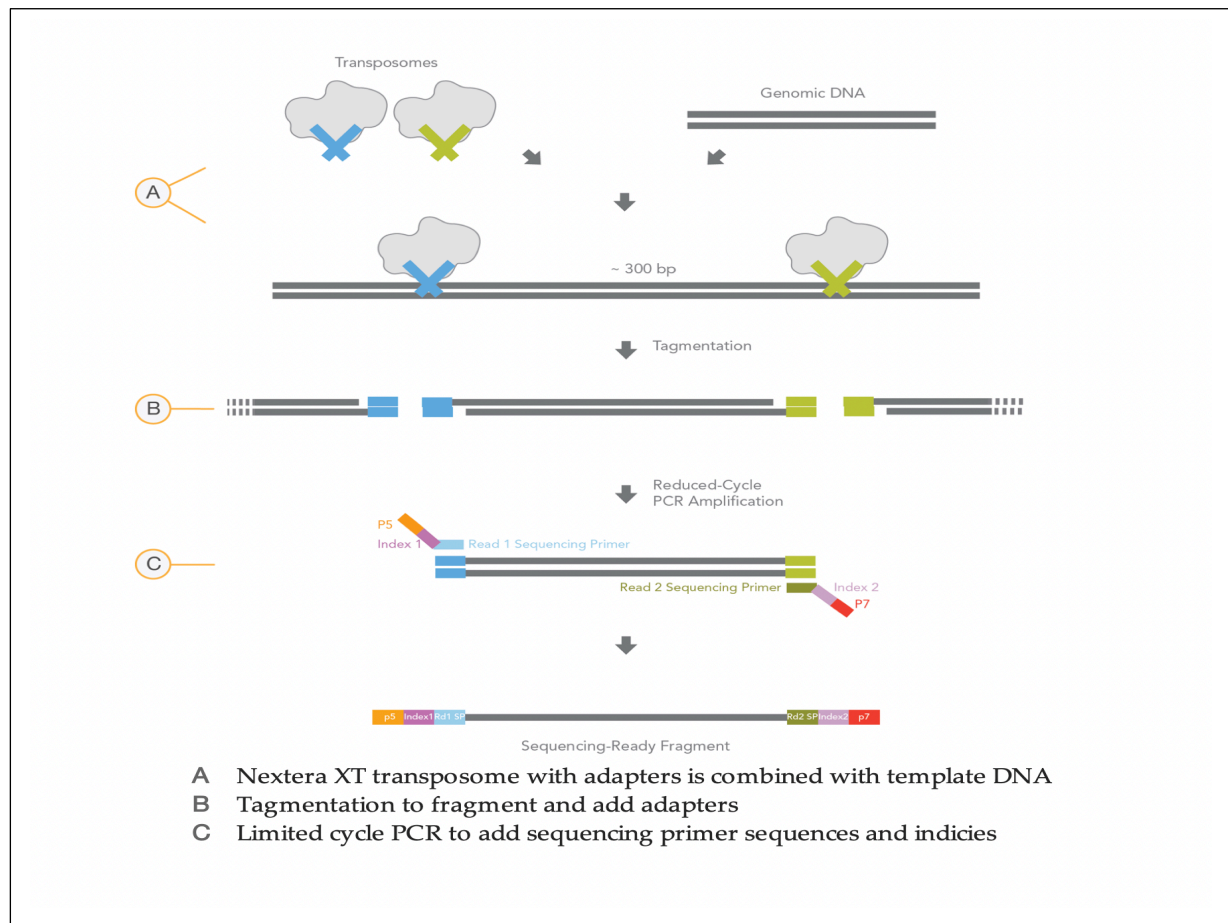


**Figure 1.4** Illustration of how the Nextera XT DNA library kit fragments and tags input DNA with unique adapter sequences.

Taken from:

[https://hahana.soest.hawaii.edu/cmoreserver/summercourse/2015/documents/Metagenomics\\_06-22/nextera\\_xt\\_sample\\_preparation\\_guide\\_15031942\\_c.pdf](https://hahana.soest.hawaii.edu/cmoreserver/summercourse/2015/documents/Metagenomics_06-22/nextera_xt_sample_preparation_guide_15031942_c.pdf). Accessed 22 March 2022.

In the second step adapter sequences are annealed to the 5' and 3' end of each of the DNA fragments or amplicons. These adapters contain known sequences, one adapter contains the sequence for the primary annealing site, whilst the other adapter is used to anchor the sequence for sequencing.

Size selection occurs in the third step to select fragments of the size compatible with the sequencing platform. Size selection can be done by gel electrophoresis or with a bead-based size selection step. With the Illumina® (San Diego, CA, USA) based DNA protocol, a bead-based system is used. Amplicon based libraries usually don't need to undergo size selection if the PCR primers used for enrichment of the gene of interest are prespecified to generate correct size amplicons.

Final library quantification and quality control are the last steps in the library preparation. Methods include use of the Bioanalyzer® system (Agilent, Santa Clara, CA, USA) which can provide information on both the fragment size of the library as well as the library quality. Our preferred method to quantify DNA is the Qubit™ (ThermoFisher, Waltham, USA) system which uses double-stranded DNA fluorescent dye which intercalates into the double-stranded DNA of the sample. The dye then becomes excited in the fluorometer, and the concentration of the DNA is determined by fluorescence intensity.

### **1.1.8.2 Sequencing:**

Illumina® (San Diego, CA, USA) sequencing chemistry will be overviewed here and consists of paired end sequencing. In fact, this is sequencing by synthesis as well as paired end sequencing. The flow cell is a glass slide with lanes each coated with a lawn on the surface containing thousands of oligonucleotides.

Following sample preparation, the first step in Illumina® (San Diego, CA, USA) sequencing is called clustering. This step allows hybridisation of the patient sample with its attached adaptor oligo to a complementary oligo on the flow cell. A polymerase then creates a complement of the hybridised fragment. Following creation of the complement strand, the double stranded molecule is denatured, and the template is washed away. Each strand is further clonally amplified by bridge amplification. This means that the strand still hybridised to the flow cell, will fold over and form a bridge, binding on the other end to its complementary oligo. A polymerase will once again form a complement strand and a double stranded bridge is formed; this bridge is denatured forming two single stranded molecules that are

each tethered to the flow cell. This process is repeated simultaneously for millions of strands bound across the lawns of the flow cell. After bridge amplification, the reverse strands are cleaved off and washed away, leaving the forward strand of which the 3' ends are blocked so that unwanted binding and amplification do not occur.

After sufficient clusters have been generated, it's time for sequencing to commence. Sequencing begins with the extension of the first sequencing primer to produce the first primer with the first read. With every cycle, fluorescently tagged nucleotides compete for addition to the complementary strand and will bind if the complement is present. After the addition of each nucleotide, the clusters are excited by a light source from the analyser, measuring a characteristic fluorescent signal that is emitted from the nucleotide that was added. This process is called sequencing by synthesis.

Sequences from each patient can be separated based on the unique sequencing adapter or index ligated to each sample during library preparation. The sequences generated are then aligned to the reference genome using bioinformatic tools which are described in more detail below. This is the method our group used for our study. The Illumina® (San Diego, CA, USA) platform was chosen for this study for cost implications and for downsizing possibilities.

## **1.2 Study aims and objectives**

Currently, we are the only laboratory in the Gauteng region (within both the public and private sectors) to perform *ABL1* sequencing in CML patients with failed or suboptimal response to TKI. The gold standard we follow is Sanger Sequencing, a method which requires a considerable amount of scientist "hands on" time" and is costly done on one or few samples at a time.

Given the increasing number of diagnostic demands and the limited number of scientists available the head of the unit initiated the design of a specific *BCR-ABL1* NGS panel to increase automation of the test and reduce cost. This initiative was further supported by reports from overseas laboratory and consortiums who

recommended NGS over SS as the former increase sensitivity to 3% versus 20% and detects early and complex mutations.

To this effect a team consisting of a pathologist, (Dr P Willem, Head of Unit), a senior scientist (Mishalan Moodley) and myself, together with the assistance of an Illumina® representative (Marija Kvas) strategized and designed an initial NGS validation study for the routine detection of *ABL1* mutations in CML patients.

Within in the frame of this study the aims of this MMED were:

1. To choose an appropriate retrospective cohort of CML patients with documented history and mutational status. To extract clinical and demographic data as well as frozen RNA samples to assess inclusion criteria regarding both patients' information and samples quality.
2. To validate various protocols of the *ABL1* NGS panel design and sequencing method on a retrospective validation cohort.
3. To compare the results to those of SS.
4. To discuss the next steps necessary for the implementation of this test in a routine diagnostic laboratory.

## 2 METHODOLOGY

---

### 2.1 Research Question

Can the *ABL1* NGS method designed in our laboratory detect mutations previously identified by SS in *ABL1* KD with an increased cost effectivity?

## **2.2 Materials and methods:**

### **2.2.1 Study Population:**

- Retrospective samples from 40 CML patients were used in this study. These samples had previously been tested in the Somatic Cell Genetics Unit (SCGU) for mutations in *ABL1* using SS, requested by the attending clinician over a 5-year period (2015 – 2020). Samples were selected from three hospitals in the greater Johannesburg area: Charlotte Maxeke Johannesburg Academic Hospital (CMJAH), Chris Hani Baragwanath Academic Hospital (CHBAH) and Steve Biko Academic Hospital (SBAH) for standardisation purposes. All SS analysis was done in the SCGU, an accredited diagnostic laboratory. Samples met the WHO criteria for the diagnosis of CML. SS had had been requested on these samples due to patients showing suboptimal response to TKI therapy or loss of molecular response on TKI therapy by RQ-PCR.
- A unique study number was assigned to each sample to deidentify samples and protect patients' confidentiality (Ethics nr: M200567, obtained for the University of Witwatersrand Human Research Ethics Committee).

### **2.2.2 Sample selection and eligibility criteria:**

The current NGS library preparation commercial package for a custom panel set on the Illumina® (San Diego, CA, USA) iSeq™ 100 sequencing system accommodates for a total of 24 library preparations (Illumina® (San Diego, CA, USA) Nextera® XT DNA library preparation kit). For this reason, 40 samples were selected with an RNA concentration of  $\geq 1\mu\text{g}$ . For the initial sequencing run 24 samples were used however the sequencing run failed. For optimisation of the assay and the second sequencing run, 4 out of the 40 samples were excluded due to inadequate cDNA concentration, and a further 24 samples were utilised after the method was optimised.

### **2.2.3 RNA extraction and cDNA synthesis:**

RNA from retrospective samples, previously extracted from peripheral blood or bone marrow samples for SS analysis and stored at  $-80^{\circ}\text{C}$ , was used for NGS. RNA extraction was performed using Chemagic 360 (Perkin Elmer® (Waltham, MA, USA)), according to manufacturer's protocol. It is an automated, bead-based DNA

extractor. Total RNA (1ug) was used to synthesise cDNA with SuperscriptIV™ (ThermoFischer, Waltham, USA). Subsequently cDNA was used for first round PCR amplification of *BCR-ABL1*.

All reagents, corresponding volumes and PCR conditions used for cDNA and PCR steps are listed in Appendix A (Table 7.1).

## 2.2.4 PCR amplification for *BCR-ABL1*:

We tested validity of these samples by repeating *BCR-ABL1* PCR amplification and selected the samples that were successful, and which showed amplification after resolving on an agarose gel. Patient characteristics of all 40 samples are shown in Table 2.1.

- Samples with RNA quality and yield of <1ug were excluded.
- Samples which failed PCR amplification were excluded for further downstream sequencing.

**Table 2.1 Patient characteristics and mutations of the samples evaluated for sequencing.**

| Unique number | Sex | Age | Mutation information from SS | RQ-PCR (IS%)*      | Assay sensitivity of RQ-PCR | Molecular Response |
|---------------|-----|-----|------------------------------|--------------------|-----------------------------|--------------------|
| 0201          | M   | 53  | T315I                        | 4.28%              | MR4.5                       | MCR                |
| 0433          | M   | 32  | E255K                        | 2.89%              | MR4.5                       | MCR                |
| 0506          | M   | 39  | Y253H/T315I                  | 17.97%             | MR4.5                       | -                  |
| 0638          | M   | 47  | T315I                        | 16.49%             | MR4                         | MCR                |
| 0728          | F   | 57  | F317L                        | 4.02%              | IE                          | MCR                |
| 0797          | F   | 60  | M351T                        | No RQ <sup>‡</sup> | -                           | -                  |
| 1071          | M   | 37  | R367Q                        | 2.33%              | MR4.5                       | MCR                |
| 1224          | F   | 31  | No mutation                  | 14.03%             | MR4                         | X                  |
| 1284          | M   | 40  | M244V                        | 33,07%             | MR4,5                       | X                  |
| 1337          | M   | 46  | L364I                        | 80%                | MR5                         | X                  |
| 1415          | F   | 79  | F359I                        | 36,94%             | MR5                         | X                  |
| 1706          | F   | 41  | F317L                        | No RQ <sup>‡</sup> | -                           | -                  |
| 1781          | M   | 29  | T315I                        | 0,38%              | IE                          | CCR                |

|      |   |    |                                  |                    |        |     |
|------|---|----|----------------------------------|--------------------|--------|-----|
| 1811 | M | 16 | No mutations                     | 5.66%              | MR5    | MCR |
| 2148 | M | 40 | p.Q252pfs*46                     | 6,70%              | MR4    | MCR |
| 2548 | M | 34 | G250E                            | 1.78%              | MR4    | MCR |
| 2770 | F | 26 | E255K                            | 19,95%             | MR4,5  | X   |
| 2848 | M | 45 | G250E                            | No RQ <sup>‡</sup> | -      | -   |
| 2989 | F | 42 | F359V                            | 19,94%             | MR4,5  | X   |
| 3038 | F | 53 | Y253H                            | 11.18%             | MR4.5  | X   |
| 3064 | M | 59 | G250E                            | 8,60%              | MR4.5# | MCR |
| 3083 | M | 35 | A380T                            | 16,76%             | MR5    | X   |
| 3165 | M | 47 | M244V                            | 8,08%              | IE     | MCR |
| 3236 | F | 28 | A287V                            | 5.37%              | MR4.5  | MCR |
| 3476 | M | 27 | K247E & V299M                    | 0.20%              | '-3.51 | CCR |
| 3541 | F | 38 | V299L                            | 7,99%              | MR4,5  | MCR |
| 3770 | M | 32 | T315I                            | 0.63%              | MR4.5  | CCR |
| 3809 | F | 43 | F359V                            | 18,24%             | MR4,5  | X   |
| 3916 | M | 27 | No mutations                     | 0.16%              | MR4.5  | CCR |
| 3929 | F | 48 | F311L                            | 20,73%             | MR4    | X   |
| 3942 | M | 43 | No mutations                     | 8.12%              | IE     | MCR |
| 3991 | M | 45 | Stop codon at aa261 (TGG to TAG) | No RQ <sup>‡</sup> | -      | -   |
| 4229 | M | 40 | Y232C                            | No RQ <sup>‡</sup> | -      | -   |
| 4258 | M | 28 | 1087_1088del                     | 0,61%              | IE     | CCR |
| 4302 | F | 41 | E255K/T315I                      | No RQ <sup>‡</sup> | -      | -   |
| 4316 | F | 64 | E255K                            | 24,09%             | MR4    | X   |
| 5038 | F | 29 | G250E                            | 15,79%             | MR5    | X   |
| 5331 | F | 61 | G250E                            | 33,61%             | MR5    | X   |
| 5346 | M | 48 | G250E                            | 0.71%              | MR4    | CCR |
| 5879 | F | 57 | G250E                            | 10,40%             | IE     | X   |

CCR: Complete Cytogenetic Response; MCR: Major Cytogenetic Response; F: Female; IE: In Evaluable; M: male; MR4: ; MR4.5: ; MR5;

X – No response

<sup>‡</sup>IS% calculated according to well established international standardised parameters (54) using a correction factor of 0.57.

<sup>‡</sup>No RQ – These samples were referred from private laboratories where the RQ-PCR is performed.

- : RQ-PCR not performed in our laboratory thus Assay sensitivity and Molecular Response not available

To select the specific *BCR-ABL1* fragment we performed the following semi-nested PCR on cDNA. A forward primer targeting BCR exon 13 (55) and a reverse primer targeting *ABL1* exon 10/11 (56) was used to amplify the *BCR-ABL1* fusion gene.

This amplification step generates a 1586bp product. The semi nested second-round PCR uses a forward primer targeting BCR exon 5 (57) and the same reverse primer from first round (Figure 2.1). These primers are used as part of routine *BCR-ABL1* amplification for SS in our laboratory.

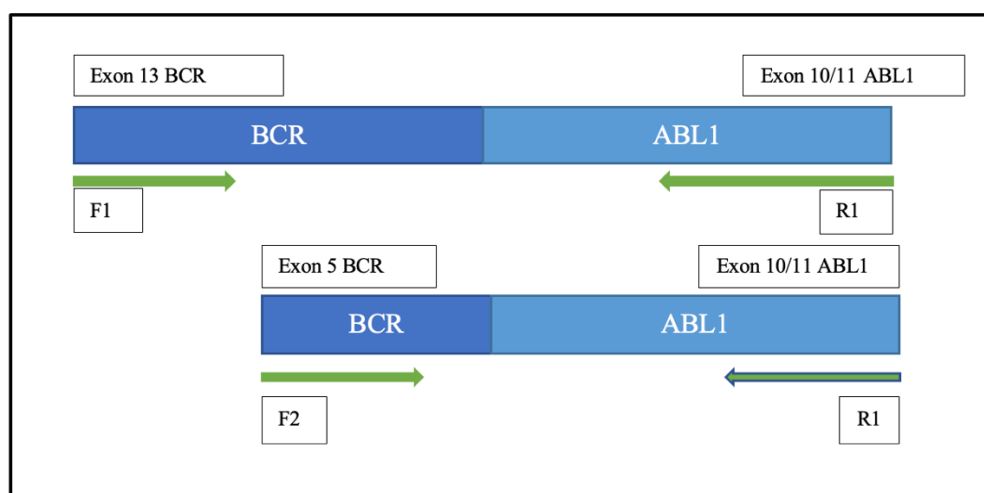


**Figure 2.1** Illustration of primers used for semi-nested PCR

### 2.2.5 Qubit assay:

To quantify DNA concentration, all samples were measured using the Qubit™ high sensitivity RNA and DNA assays. The Qubit™ analyser is sensitive in DNA quantification as Qubit™ measures DNA concentration directly. Two standards were prepared and run on the Qubit fluorometer before patient samples were analysed. Specific dsDNA standards (low and high) are included in the kit. Standards were prepared by mixing 189uL of buffer, with 1uL of fluorogenic dye and 10uL of standard. The solution was vortexed and left to incubate for 2 minutes. After 2 minutes the standards were analysed with the Qubit™ fluorometer (ThermoFisher, Waltham, USA).

To assess the correct size of the fragment and ensure amplification occurred, both first and second round PCR were resolved on a 2% Agarose gel electrophoresis (Refer to Appendix B). Based on successful amplification of the semi nested PCR,

the product from the first round PCR was used for NGS. Since the first round PCR product could not be visualised on gel electrophoresis, second round product was used to verify and proceed with sequencing of first round PCR product.

#### **2.2.6 DNA library preparation using Illumina Nextera™ XT DNA library Preparation Kit (Illumina, CA, USA):**

**Materials and equipment:** (See *Appendix C*)

##### **2.2.6.1 Nextera™ library preparation:**

As previously mentioned the Nextera™ XT DNA library preparation kit was chosen as the transposase technology used allows for DNA to be cut into smaller fragments whilst at the same time, adapters are ligated onto the ends of the fragments. This chemistry allows for quicker library preparation times as well as less hands on scientist time.

All steps for library preparation, tagmentation, and sequencing were performed exactly according to manufacturer's protocol and guidelines (Appendix D). A brief overview of all steps are shown below and in Figure 2.2.

##### **2.2.6.2 Tagmenting of genomic DNA:**

Nextera™ XT DNA Library Preparation kit makes use of transposome technology to fragment and at the same time add special adapter sequences. The reactions are carried out in a 96-well PCR plate with a sample volume calculated to contain 10ng of cDNA.

##### **2.2.6.3 Amplification of DNA libraries:**

During library amplification the tagmented DNA was amplified and the indexes required for cluster generation (i7 and i5) added to each amplicon.

Adapter sequences were allocated to each patient sample in the following manner before being added to the wells in the plate.

**Table 2.2 Sample number with assigned unique sequence adapter.**

| <b>Sample number:</b> | <b>Adapter Sequence:</b> | <b>Sample number:</b> | <b>Adapter Sequence:</b> |
|-----------------------|--------------------------|-----------------------|--------------------------|
| 0433                  | 705/503                  | 3038                  | 710/503                  |
| 0506                  | 705/505                  | 3083                  | 710/505                  |
| 0638                  | 705/506                  | 3165                  | 710/506                  |
| 0728                  | 705/517                  | 3236                  | 710/517                  |
| 0797                  | 706/503                  | 3541                  | 711/503                  |
| 1224                  | 706/505                  | 3770                  | 711/505                  |
| 1337                  | 706/506                  | 3809                  | 711/506                  |
| 1415                  | 706/517                  | 3916                  | 711/517                  |
| 1781                  | 707/503                  | 3942                  | 714/503                  |
| 1811                  | 707/505                  | 4258                  | 714/505                  |
| 2148                  | 707/506                  | 4316                  | 714/506                  |
| 2548                  | 707/517                  | 5038                  | 714/517                  |

#### **2.2.6.4 Library clean up:**

In this step libraries were be cleaned-up. This step makes use of single-sided bead purification purify the amplified libraries. Single-sided bead-based purification allows for the small fragments that were generated by the transposomes to be removed from the sample whilst the fragments of interest for sequencing remain behind in the supernatant.

#### **2.2.6.5 Normalise libraries:**

In this step the quantity of each library made with Nextera XT index kits was normalised to ensure equal representation of each library in the pooled library. We made use of standard normalisation. After running samples on the Qubit the nM concentration of each sample was calculated using the following formula:  $nM = ((ng/ul) / (avg\ bp\ size \times 660\ g/mol)) \times 1,000,000$ . Libraries were diluted 1nM which equates to 100pM final loading concentration of each sample.

#### 2.2.6.6 Library pooling:

For library pooling, 10uL of each diluted library was then added to a single tube to create a multiplexed library pool, ready for loading onto the sequencing cartridge. Each sample was quantified before pooling the libraries using the Qubit fluorometer, using the method described above. The 24 samples that were used for library preparations were quantified, only 20 of these samples were eligible for sequencing due to 4 samples not having a high enough cDNA concentration for sequencing. The pooled libraries were then loaded onto the sequencing cartridge by pipetting 20uL of the pooled library to the bottom of the reservoir of the cartridge and loading the cartridge and the flow cell into the iSeq100 sequencing system.

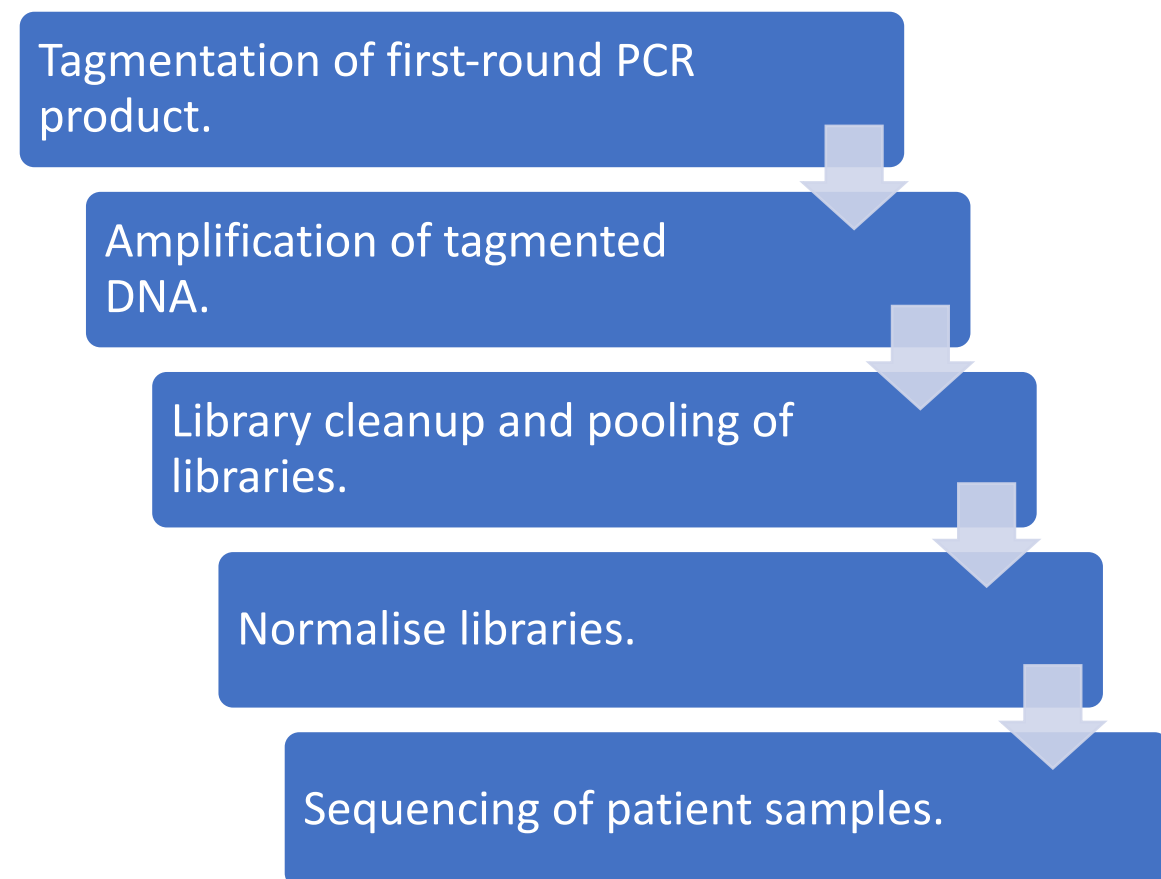


Figure 2.2 Diagram indicating the steps for library preparation and sequencing.

#### 2.2.7 Sequencing on the iSeq™ 100 system:

The analyser performs three main steps during the sequencing run. These steps are cluster generation, sequencing and analysis. Cluster generation involves the

denaturing of the library into single strands. During the cluster generation process, single DNA molecules are bound to the surface of a flow cell and are systematically amplified to form cluster.

During sequencing with every cycle, fluorescently tagged nucleotides compete for incorporation based on complementarity of the target strand. After the addition of each nucleotide, the flow cell is excited by a light source from the analyser, measuring fluorescent signals emitted from the nucleotide that was incorporated. This process is called sequencing by synthesis. The number of cycles determines the read depth and the emission wavelength with the signal intensity determines the base call. At the end of the sequence of the first read, the read product is washed away. The 3' ends of the sequences which were initially blocked, are unblocked, the template strand once again folds over to form a bridge. Polymerases extend the second flow cell and adds in the complementary nucleotides. The bridge is again denatured and this time the forward strand is washed away leaving the reverse strand behind for sequencing. Fluorescently labelled nucleotides are once again incorporated, and the fluorescent intensity measured to determine the base call.

During sequencing images of clusters are taken, which with the iSeq™ 100 technology, involves one-dye chemistry using a one fluorescent label and two imaging cycles to encode the data of one of the four nucleotides binding to the strand. The first imaging cycle detects adenine (A) and thymine (T). A chemical process then cleaves the dye from adenine and simultaneously adds a similar dye to cytosine (C). The second imaging cycle detects cytosine and thymine. In real-time analysis software will then perform base calling, filtering and quality scoring of the sequencing process.

### **2.2.8 Analysis:**

As the sequencing run progresses, control software automatically transfers base call files to a pre-named output folder, stored on the analyser, for data analysis.

## **2.3 Data analysis:**

NGS generates vast amounts of data that need to be reorganised into relevant and understandable information. For this phase of analysis scientists use various applied bioinformatic software.

### **2.3.1 Quality control of sequencing data:**

During the sequencing process, certainty of base calling is important to ensure accuracy of the base inserted. A higher Q-score implies that the accuracy of a base call is of good quality and therefore more likely to be correct. This quality scoring system is based on a modified version of the Phred algorithm (58), which is an algorithm that assigns a quality value to bases to determine the accuracy of the base called during the sequencing read. With Illumina sequencing technology, quality scoring is performed by calculating a set of predictors for each base call. These predictor values are then used to look up the Q-score in a quality Table. A Q-score of 30 or more is defined as a good Q-score and the probability for erroneous base calling is 1 in 1000. In general, NGS experiments use a Q-score  $\geq 30$  for further downstream analysis. The Trimmomatic software (59) tool was used to trim all adapters and retain bases with a Phred score  $\geq 30$ .

### **2.3.2 Mapping and variant calling:**

After QC of NGS, three main steps are used for NGS analysis: alignment, variant calling and analysis. Analysis of sequencing runs are carried out by aligning sequencing reads to the reference genome: either to the gene (*ABL1*) or to chromosome 9 (Hg38). The variant caller software (Illumina®) separates samples' sequencing reads using each barcode. Sequencing reads were aligned to chromosome 9 using Burrows-Wheeler-Aligner (BWA). BWA is a software package that is used for mapping of low-divergent sequences against a large reference genome. The resulting output Binary Alignment Files (BAM) files are used to generate a Variant Call Format (VCF) file which only contains mutations. Figure 2.3 illustrates the pipeline used to generate the final VCF file.

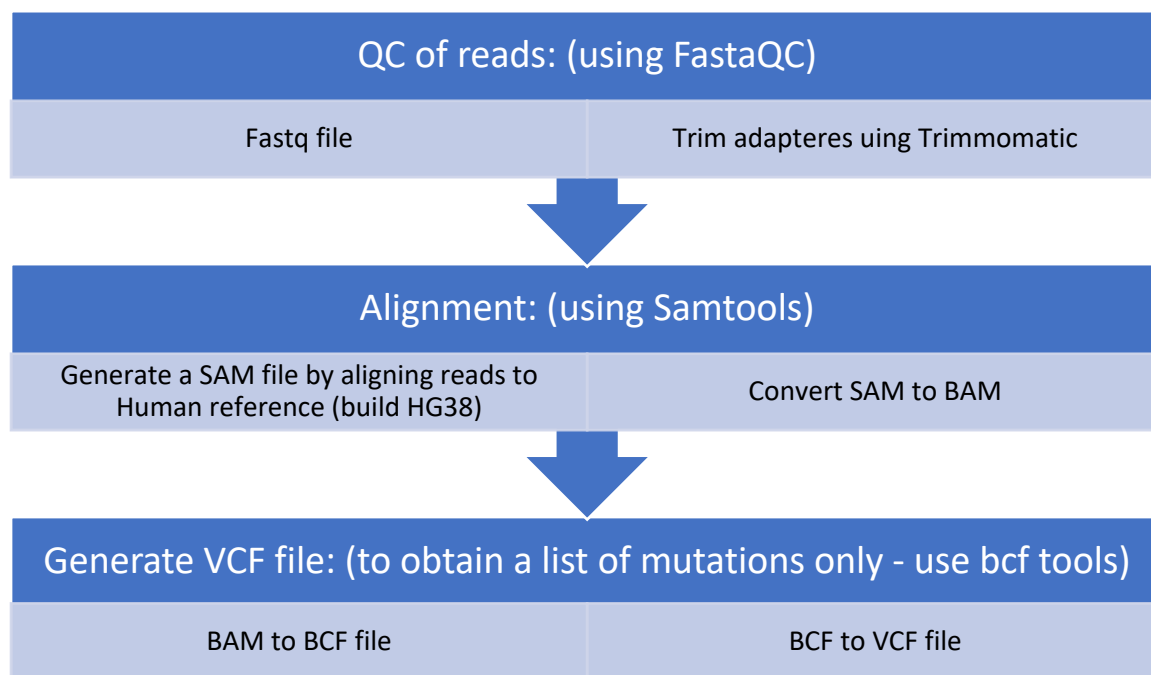


**Figure 2.3 Bioinformatic steps used for analysis.**

QC: quality control; SAM: Sequencing Alignment Map; BAM: Binary Alignment Map; BCF: Binary Variant Call Format; VCF: Variant Call Format.

Our first attempt to sequence the first PCR round product as the *BCR-ABL1* target using the routine primers essentially led to a failed sequencing run (see results section). In effect, the alignment of reads revealed non-specific sequencing of multiple targets throughout the genome. Based on this result and to confirm the source of failure the first-round product was also resolved on the LabChip® GX Touch (Perkin Elmer® (Waltham, MA, USA)) to visualise the first-round amplification product of BCR-ABL. The LabChip® GX Touch is a capillary gel electrophoresis analyser that is more sensitive than agarose gel electrophoresis in detecting and resolving products from PCR.

In addition to non-specific amplification seen on the agarose gel electrophoresis, the routine set of primers showed non-specific amplification as evidenced by additional bands as well as primer overload based on primer-dimer formation noted. A new set of primers were ordered, and steps taken are described below.

### **2.3.3 Second attempt at *BCR-ABL1* amplification with published primers:**

For the second attempt at *BCR-ABL1* amplification and to ensure more specific target amplification, a new set of first round primers were used to improve the quality and specificity of the targeted region. Primers originally designed by Soverini *et al.* were used (50). In fact, these are primers they use for routine amplification of the full *BCR-ABL1* product. The forward primer targets BCR exon 12-13 (5' - GAGCAGCAGAAGAAGTGTTTCAGS-3'), the reverse primer is located on *ABL1*, exon 10 (5'-CTTGGAGTGAGGCATCTCAG-3').

The PCR product was purified using the ExoSAP-IT™ Express reagent. This process removes any unbound nucleotides and primers. 4uL of ExoSAP-IT™ was added to each sample, samples were then placed in the thermocycler with the following conditions: 37°C for 4 minutes on one cycle which degrades excess primers and nucleotides, then 80°C of 1 minute for 1 cycle to inactive the ExoSAP-IT™ Express reagent and hold at 4°C. The reaction occurs in the thermocycler with a heated lid. PCR product concentration was quantified using Qubit™ fluorometer (ThermoFisher, Waltham, USA) and the Qubit™ dsDNA HS (high sensitivity) kit (ThermoFisher, Waltham, USA) as described above.

Twenty-four samples were subsequently selected from the PCR reactions to undergo library preparation and sequencing. These samples were selected based on the gel electrophoresis that showed clear bands indicating amplification had taken place. The same methods described above for library preparation and sequencing were followed.

## **3 RESULTS:**

---

### **3.1 First attempt using the first PCR amplification product to sequence *ABL1*:**

#### **3.1.1 Gel electrophoresis of initial first and second round PCR amplification:**

With the use of our routine *BCR-ABL1* primers, the aim was to sequence the first-round fragment of 1586bp. Although first round cannot be visualised on agarose gel

electrophoresis, a second round PCR step was performed in order to confirm first round amplification using a 2% agarose gel electrophoresis (Figure 3.1).

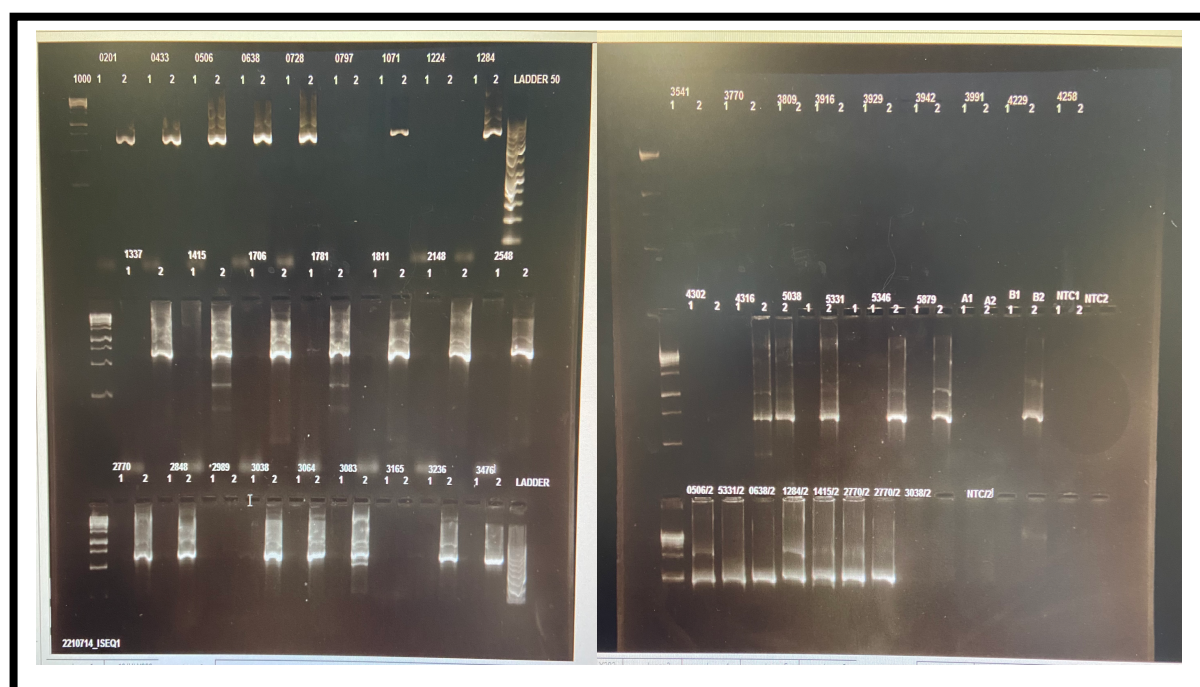


**Figure 3.1 Results of 2% Agarose gel electrophoresis.**

Samples were loaded onto the gel in order of first round first (labelled 1) and second round, second (labelled 2). Above each pair the unique sample identification number is written. As noted, the wells labelled 1 no bands are noted, this is the first round PCR product which cannot be visualised on the gel electrophoresis. Where the wells are labelled 2 (indicating 2<sup>nd</sup> round PCR), majority of samples show a band of the correct 593bp product, but also spurious bands (>593bp) are noted which indicates non-specific amplification. The ladder is indicated on the far left and ranges from 250 bp through to 1000 bp in size. Non-template controls (NTC's) are indicated in the image on the right and show no amplification indicating that contamination is absent.

### 3.1.2 Results of patient samples after quantification and clean-up:

After PCR, samples were purified to get rid of unwanted enzymes, nucleotides, primers and other components that may interfere with other downstream processes. Cleaned-up PCR product was measured with the Qubit assay (Table 3.1).

**Table 3.1 cDNA concentration, as measured on Qubit™ fluorometer, (ThermoFisher, Waltham, USA) before and after clean-up.**

| Unique number | Concentration of cDNA before clean up (ng/uL) | Concentration of cDNA after clean up (ng/uL) | Total concentration (ng/uL) in final volume of 18uL) |
|---------------|-----------------------------------------------|----------------------------------------------|------------------------------------------------------|
| 0201          | 7.10                                          | 0.80                                         | 14.40                                                |
| 0433          | 5.64                                          | 0.90                                         | 16.20                                                |
| 0506          | 5,88                                          | 1.44                                         | 25.92                                                |
| 0638          | 11.00                                         | 2.04                                         | 36.72                                                |

|      |      |      |       |
|------|------|------|-------|
| 0728 | 6.46 | 1.09 | 19.62 |
| 0797 | 5.46 | 1.24 | 22.32 |
| 1071 | 5.87 | 0.88 | 15.84 |
| 1224 | 7.25 | 1.51 | 27.18 |
| 1284 | 4.90 | 0.72 | 12.96 |
| 1337 | 5.11 | 1.07 | 19.26 |
| 1415 | 9.82 | 1.66 | 29.88 |
| 1706 | 5.74 | 0.68 | 12.24 |
| 1781 | 6.78 | 0.99 | 17.82 |
| 1811 | 5.81 | 0.10 | 1.80  |
| 2148 | 5.57 | 0.70 | 12.60 |
| 2548 | 8.08 | 0.92 | 16.56 |
| 2770 | 5.78 | 0.76 | 13.68 |
| 2848 | 6.50 | 0.85 | 15.30 |
| 3038 | 7.59 | 1.33 | 23.94 |
| 3064 | 5.41 | 0.74 | 13.32 |
| 3083 | 7.26 | 1.08 | 19.44 |
| 3165 | 4.97 | 0.93 | 16.74 |
| 3236 | 6.75 | 1.70 | 30.60 |
| 3476 | 5.01 | 0.81 | 14.58 |
| 3541 | 6.51 | 1.18 | 21.24 |
| 3770 | 5.69 | 1.23 | 22.14 |
| 3809 | 5.42 | 1.14 | 20.52 |
| 3916 | 6.65 | 1.00 | 18.00 |
| 3929 | 6.69 | 0.86 | 15.48 |
| 3942 | 6.83 | 1.13 | 20.34 |
| 3991 | 6.13 | 0.78 | 14.04 |
| 4229 | 4.75 | 0.74 | 13.32 |
| 4258 | 6.06 | 0.99 | 17.82 |
| 4302 | 6.80 | 0.79 | 14.22 |
| 4316 | 9.53 | 2.12 | 38.16 |
| 5038 | 7.17 | 0.93 | 16.74 |
| 5331 | 5.59 | 1.12 | 20.16 |
| 5346 | 6.39 | 1.39 | 25.02 |
| 5879 | 5.87 | 0.74 | 13.32 |

After purification the concentration of PCR product decreases significantly. Samples were quantified again. Samples  $\geq 10\text{ng/uL}$  were included in the run, which left 20

sample eligible for library preparation. These samples together with Qubit concentration values are listed in Table 3.2.

**Table 3.2 List of samples used for NGS processing (first sequencing run).**

| Unique number | Concentration of cDNA before clean up (ng/uL) | Concentration of cDNA after clean up (ng/uL) | cDNA quantity after library equalising (ng/uL) |
|---------------|-----------------------------------------------|----------------------------------------------|------------------------------------------------|
| 0433          | 5.64                                          | 0.90                                         | 0.32                                           |
| 0506          | 5,88                                          | 1.44                                         | 0.61                                           |
| 0638          | 11.00                                         | 2.04                                         | 1.08                                           |
| 0728          | 6.46                                          | 1.09                                         | 0.41                                           |
| 0797          | 5.46                                          | 1.24                                         | 0.36                                           |
| 1337          | 5.11                                          | 1.07                                         | 0.59                                           |
| 1415          | 9.82                                          | 1.66                                         | 1.30                                           |
| 1781          | 6.78                                          | 0.99                                         | 0.27                                           |
| 1811          | 5.81                                          | 0.10                                         | 0.46                                           |
| 2548          | 8.08                                          | 0,92                                         | 0.62                                           |
| 3038          | 7.59                                          | 1.33                                         | 0.30                                           |
| 3165          | 4.97                                          | 0.93                                         | 0.49                                           |
| 3541          | 6.51                                          | 1.18                                         | 0.23                                           |
| 3770          | 5.69                                          | 1.23                                         | 0.25                                           |
| 3809          | 5.42                                          | 1.14                                         | 0.28                                           |
| 3916          | 6.65                                          | 1.00                                         | 0.48                                           |
| 3942          | 6.83                                          | 1.13                                         | 0.31                                           |
| 4258          | 6.06                                          | 0.99                                         | 0.28                                           |
| 4316          | 9.53                                          | 2.12                                         | 1.13                                           |
| 5038          | 7.17                                          | 0.93                                         | 0.43                                           |

### 3.1.3 Sequencing results and analysis:

Below follows the results of the sequencing analysis and results.

#### 3.1.3.1 QC of sequencing reads:

The iSeq100 Sequencing System generates raw Fastq files which contain all sequencing reads. The QC for each sample was checked using FastQC software. Figures 3.2 and 3.3 represent QC reads of a single sample.

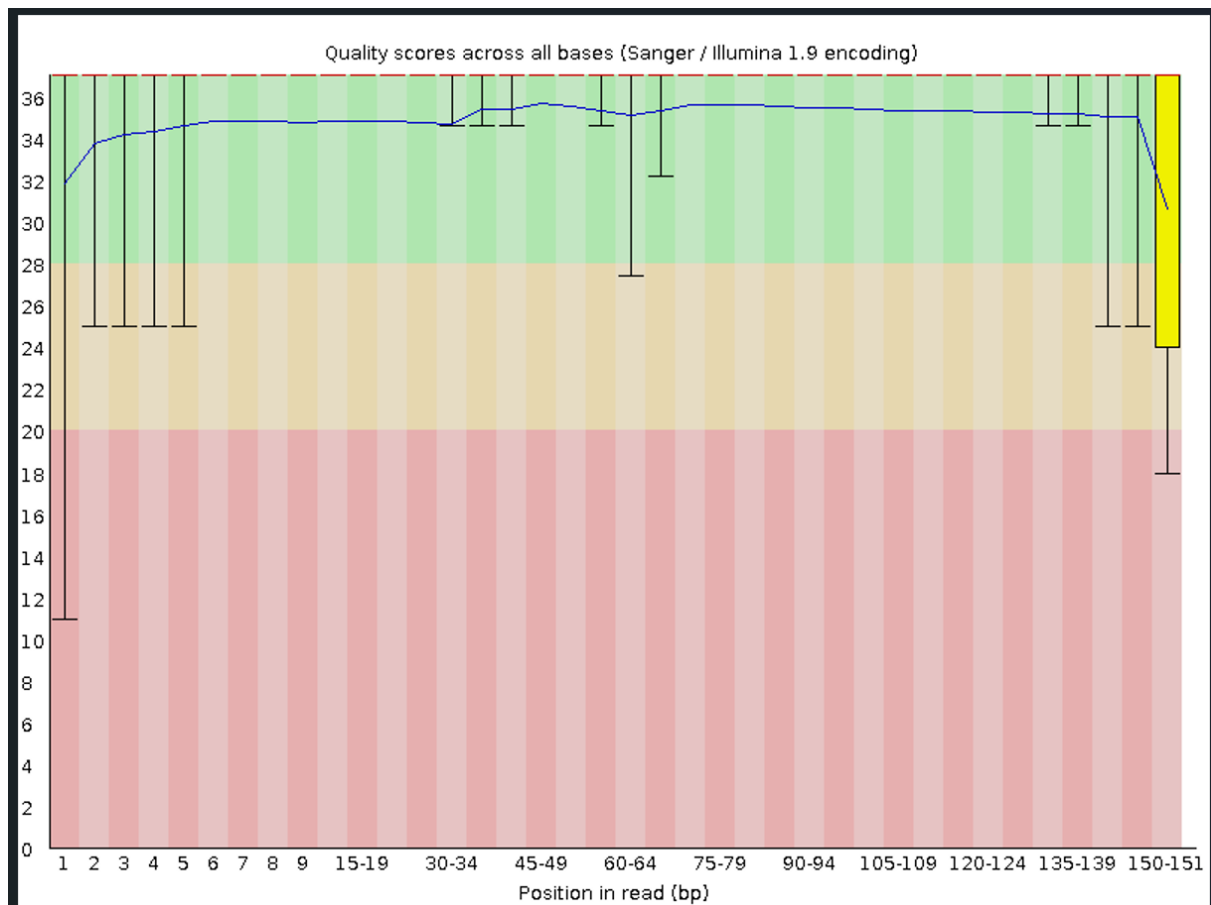


**Figure 3.2 Average of the Quality Scores for the sequencing run.**

The average quality scores are noted to be above 30 for all reads as indicated on the y-axis. The x-axis indicates the position of bases sequenced. As mentioned previously the amplicon size generated during library preparation is around 150bps.

The QC also showed that no adapters were present in all reads (Figure 3.3).

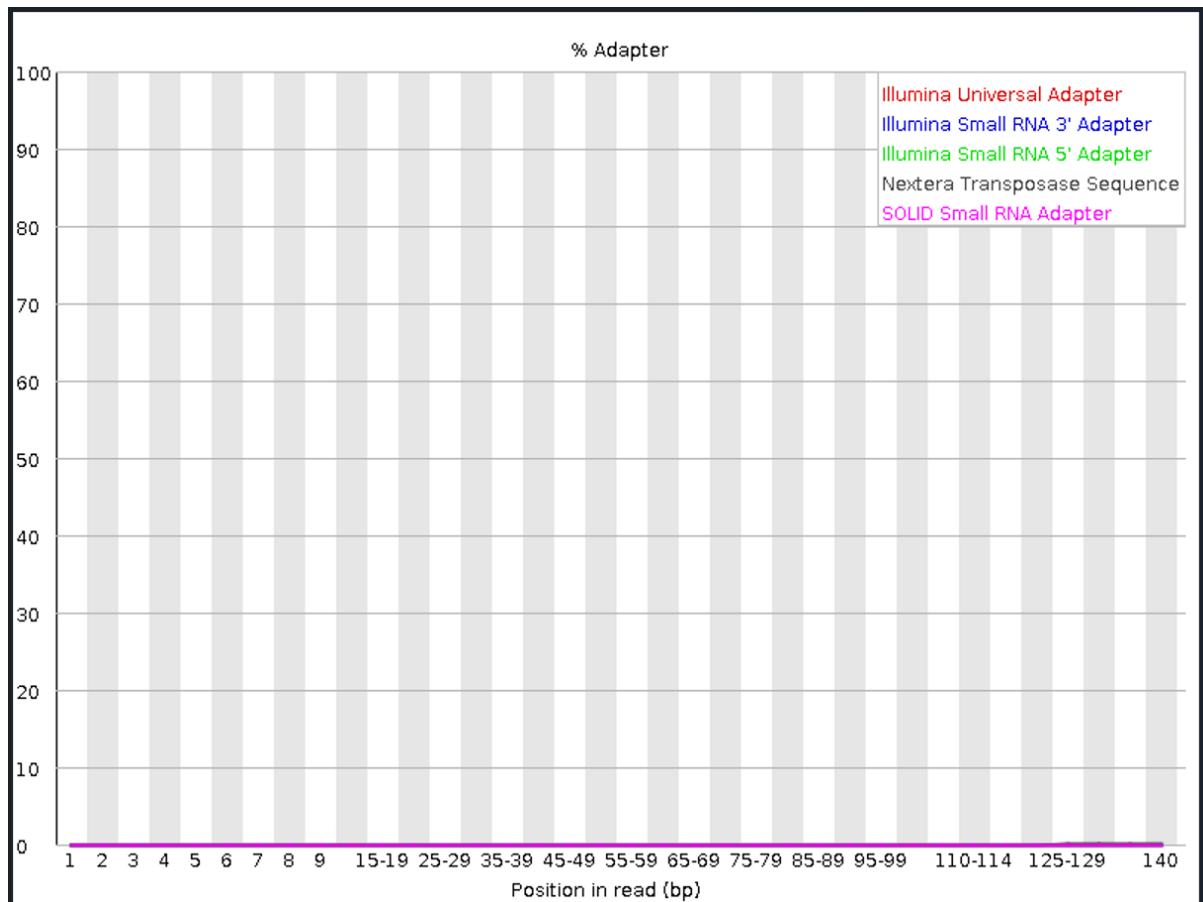


**Figure 3.3 Indicates the absence of sequencing of sequence adaptors during the sequencing run.**

Indicates whether any adapters were sequenced during the sequencing run. No adapters were sequenced during this run.

### 3.1.3.2 Trimming, aligning and variant calling:

Trimming of adapters and sequencing bases with poor Phred scores (<33) was carried out with the Trimmomatic software tool mutations (59).

### 3.1.3.3 First attempt NGS sequencing of first round *BCR-ABL1* fragment using routine in-house PCR primers:

Analysis of the VCF file, several variants with low depth of coverage (4x-10x) were detected. These variants belonged to several genes across the entire genome, not only *ABL1*. Since the few variants that were detected in *ABL1* were of extremely low coverage (4x) for all, this deemed the sequencing run to be unsuccessful, as in diagnostic sequencing of malignancies a minimum coverage of 500x is required. A snapshot of a VCF file for one patient can be found in Appendix E.

To confirm the specificity of first-round PCR, first-round product was also resolved on the more sensitive capillary gel electrophoresis, LabChip® GX Touch (Perkin Elmer® Waltham, MA, USA)). Results are illustrated in Figure 3.4.

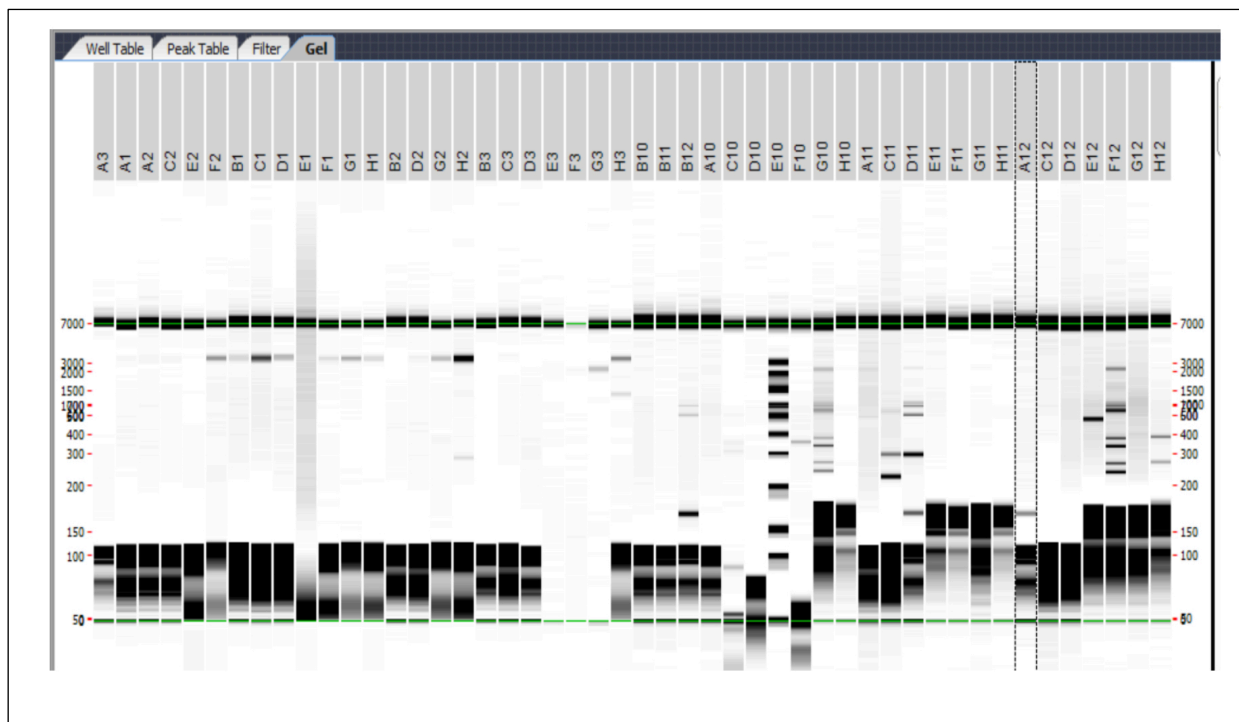


**Figure 3.4 Results of samples run on the LabChip® GX Touch.**

Annotations at the top of the Figure represents a row (A-H) and well number (1-8), that corresponds with the 96-well plate. Each sample was assigned a row and number prior to loading onto the plate. The ladder, indicating bands of different sizes can be seen on the left and right of the Figure ranging from 50 – 7000bps.

Analysis of the sample using the LabChip® GX Touch showed a consistent band around the 3000bp fragment size. This size however does not correspond with the 1585bp fragment that should be amplified with first round PCR amplification. Of note are the numerous spurious bands noted across all samples in the 50-150bp range as well as additional bands noted in some of the other samples. The bands noted in the 50-150bp range are favoured to be primer dimers that formed during amplification due to excess primers that are not participating in the *BCR-ABL1* amplification reaction. The consistent band at the 3000bp fragment size is favoured to be the first round *BCR-ABL1* product. After communication with the suppliers' technical staff (Perkin Elmer® (Waltham, MA, USA)), this size discrepancy may be due to inference with size fragment analysis from the multiple primer dimers. In addition, the ladder

was not within the normal size limit and the kit was suspected to be faulty due to interference from shipping.

Because the first round PCR fragment for sequencing failed to yield results and because of the presence of bands of the incorrect size as well as other multiple bands, indicating non-specific amplification during PCR reactions, a decision was made to sequence the second round PCR product using a different set of PCR primers.

### 3.2 Resequencing of *ABL1* using a different set of published primers and the product of the semi nested *BCR-ABL1* product as target:

In addition to sequencing the semi-nested/second round PCR product, a new set of primers (50) were also used since these primers were previously successfully used to enrich for *BCR-ABL1* sequencing in CML patients (Figure 3.5). The rationale behind this decision was that the second semi-nested PCR product is more specific than the first-round product.

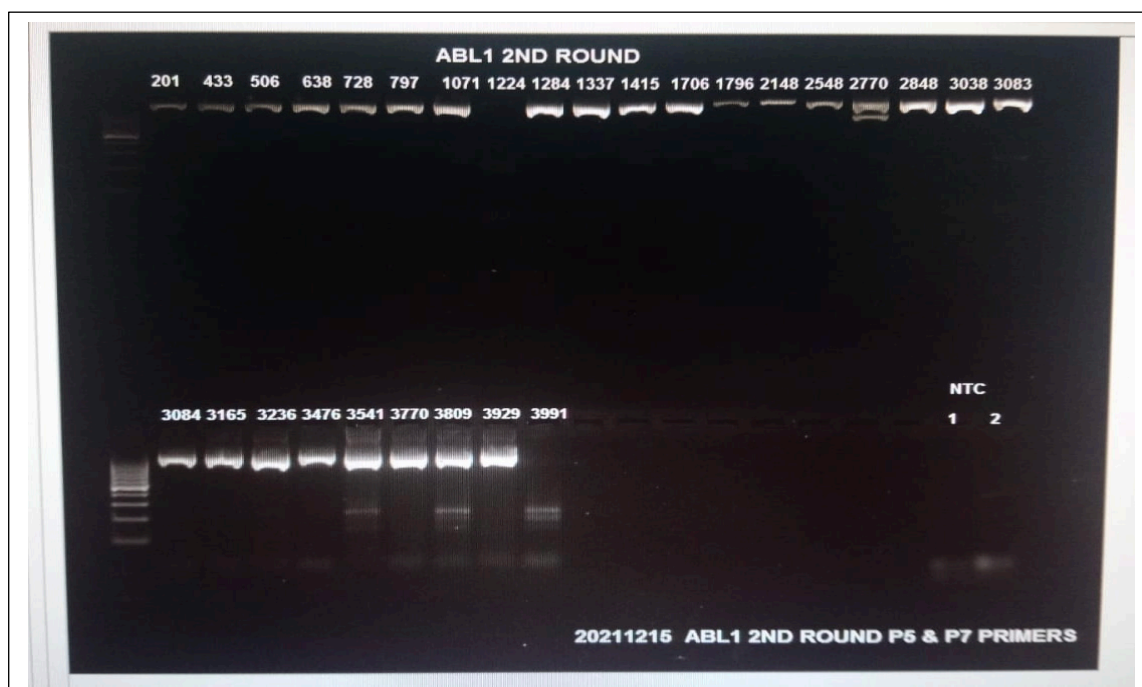


**Figure 3.5 Results of second round PCR using Soverini primers.**

On the far right is the ladder for each row of samples. The ladder is included ranging from 250bp to 1000bp. At the top of each sample, the lane is annotated with the unique number assigned for each sample. All samples show successful amplification except for sample 1224. Non-template controls (NTC) are included and show no amplification.

PCR products were cleaned-up, tagmented and quantified as described above (section 5.3.4). The concentrations and volume of each sample after clean-up is shown in Table 3.3.

**Table 3.3 Concentration of second round product for sequencing.**

| <b>Sample</b> | <b>Concentration (ng/uL)</b> | <b>Final DNA concentration (ng) in 25uL</b> |
|---------------|------------------------------|---------------------------------------------|
| 0201          | 36.7                         | 917ng                                       |
| 0433          | 59                           | 1475ng                                      |
| 0506          | 14.7                         | 367ng                                       |
| 0638          | 34.9                         | 872ng                                       |
| 0728          | 60                           | >1500ng                                     |
| 0797          | 48.5                         | 1212.5ng                                    |
| 1071          | 24.9                         | 622.5ng                                     |
| 1284          | 52                           | 1300ng                                      |
| 1337          | 60                           | >1500ng                                     |
| 1415          | 54                           | 1350ng                                      |
| 1706          | 27.8                         | 695ng                                       |
| 1781          | 26.5                         | 662ng                                       |
| 2148          | 55                           | 1375ng                                      |
| 2548          | 58                           | 1450ng                                      |
| 2770          | 60                           | >1500ng                                     |
| 2848          | 46.2                         | 1150ng                                      |
| 3038          | 44.1                         | 1100ng                                      |
| 3083          | 44.4                         | 1180ng                                      |
| 3084          | 47.2                         | 1180ng                                      |
| 3165          | 29.6                         | 740ng                                       |
| 3236          | 52                           | 1300ng                                      |
| 3476          | 23.4                         | 585ng                                       |
| 3541          | 53                           | 1325ng                                      |
| 3770          | 59                           | 1475ng                                      |
| 3809          | 36                           | 900ng                                       |
| 3929          | 52                           | 1300ng                                      |

The same bioinformatics pipeline was used as describe above (section 5.4.2). QC of NGS showed that most samples had good quality reads.

### 3.2.1 Results of second sequencing run:

Table 3.4 Concordance between NGS and variants previously confirmed on Sanger sequencing.

| Unique number | Mutation on SS: | RQ-PCR (IS%) | Assay sensitivity (log) | NGS result    | Depth of coverage (reads) mRNA analysis (NM_005157.6) | Reference Allele | Variant Allele | Variant allele frequency (%) |
|---------------|-----------------|--------------|-------------------------|---------------|-------------------------------------------------------|------------------|----------------|------------------------------|
| 0201/18       | T315I           | 4.28%        | MR4.5                   | T315I         | 5690                                                  | 185              | 5505           | 96,75%                       |
| 0433/17       | E255K           | 2.89%        | MR4.5                   | E255K         | 6580                                                  | 83               | 6497           | 98,74%                       |
| 0506/16       | Y253H/T315I     | 17.97%       | MR4.5                   | Y253H + T315I | 5153 + 5891                                           | 3825 + 2438      | 1328 + 3453    | 25,77% + 58,61%              |
| 0638/20       | T315I           | 16.49%       | MR4                     | T315I         | 5938                                                  | 63               | 5875           | 99,94%                       |
| 0728/18       | F317L           | 4.02%        | IE                      | F317L         | 6044                                                  | 14               | 6030           | 99,77%                       |
| 1071/17       | R367Q           | 2.33%        | MR4.5                   | R367Q         | DP<130                                                |                  |                |                              |
| 1284/19       | M244V           | 33,07%       | MR4.5                   | M244V         | 5762                                                  | 680              | 5082           | 88.20%                       |
| 1337/20       | No mutation     | 80%          | MR5                     | No Mutation   | -                                                     | -                | -              | -                            |
| 1415/20       | F359I           | 36,94%       | MR5                     | F359I         | 5984                                                  | 14               | 5970           | 99.77%                       |
| 1781/20       | T315I           | 0,38%        | IE                      | T315I         | 5936                                                  | 4345             | 1591           | 26,80%                       |
| 2148/19       | p.Q252pfs*46    | 6,70%        | MR4                     | No mutation   | -                                                     | -                | -              | -                            |
| 2548/20       | G250E           | 1.78%        | MR4                     | G250E         | 4347                                                  | 59               | 4288           | 98,64%                       |
| 2770/19       | E255K           | 19,95%       | MR4.5                   | E255K         | 5443                                                  | 2818             | 2625           | 48,23%                       |
| 2848/20       | G250E           | No RQ        | -                       | G250E         | 3884                                                  | 115              | 3769           | 97,04%                       |
| 3038/18       | Y253H           | 11.18%       | MR4.5                   | Y253H         | 6252                                                  | 44               | 6208           | 99,30%                       |
| 3064/20       | G250E           | 8,60%        | MR4.5                   | G250E         | 6075                                                  | 1776             | 4299           | 70,77                        |
| 3083/19       | A380T*          | 16,76%       | MR5                     | A380T         | 11010                                                 | 11 000           | 10             | 0,09%                        |

|         |        |        |       |       |      |      |      |        |
|---------|--------|--------|-------|-------|------|------|------|--------|
| 3165/19 | M244V  | 8,08%  | IE    | M244V | 5307 | 1443 | 3864 | 72,81% |
| 3236/18 | A287V  | 5.37%  | MR4.5 | A287V | 5921 | 5161 | 760  | 12,83% |
| 3476/15 | K247E* | 0.20%  | -3.51 | K257E | 6534 | 6500 | 34   | 0.96%  |
| 3541/18 | V299L  | 7,99%  | MR4.5 | V299L | 5177 | 16   | 5161 | 99,69% |
| 3770/17 | T315I  | 0.63%  | MR4.5 | T315I | 5926 | 1975 | 3951 | 66,67% |
| 3809/19 | F359V  | 18,24% | MR4.5 | F359V | 5780 | 33   | 5747 | 99,43% |
| 3929/18 | F311L  | 20,73% | MR4   | F311L | 5263 | 710  | 4553 | 86,51% |

\* Mutations that could not be seen in the VCF file but were noted to be present on analysis of the raw data.

IS%: *BCR-ABL1* expressed as a ratio to GUSB

Next generation sequencing utilising the Illumina® iSeq™ 100 Sequencing system was able to accurately identify 23 out of 24 mutations originally detected by Sanger sequencing. For samples 1071, 3083 and 3476 the mutations were not seen in the VCF files, however on review of the BAM files, the sequences were picked up by NGS and are postulated to not have been recorded in the VCF file due to the number of reads being too low for the mutations (<2%) to be validated.

Essentially the aim was to see if NGS could detect all the mutations initially detected by SS. Our results concur with SS thus proving that this NGS method can be used to detect *ABL1* mutations. Sample 1337/20 showed no detectable mutation on SS. This sample was specifically included to assess whether NGS may detect any mutations not initially found on SS. NGS does however concur with SS in this case with no additional mutations detected.

#### **4 DISCUSSION:**

---

We set out to evaluate the ability of the Illumina® iSeq™ 100 Sequencing System for the detection of *ABL1* KD domain mutations in CML patients with known mutations, previously detected with SS. Switching from SS to NGS was prompted by multiple previous retrospective studies assessing for the presence of clinically relevant low-level mutations in patients who showed “warning” and “failure” responses during the course of TKI therapy and by the need to process several samples in a single assay given the paucity of scientists in our unit. In addition, the 2020 CML guidelines from both the ELN and NCCN now recommend utilising NGS for ABL KD mutation analysis as this method is proven to be more sensitive in detecting low-level mutations (16).

As previously mentioned, the SCGU at the University of Witwatersrand is the only laboratory in the Gauteng region which provides *ABL1* KD mutational analysis for both the private and state sector. In addition, the unit receives samples from numerous centres outside the province. Owing to the availability of more generations of TKI's in South Africa (SA), the increasing diagnostic demand, the limited number of scientists, and the cost of SS, a more cost effective (see Table 1.2) and less

labour-intensive method is required for a high throughput laboratory such as the SCGU. The drawbacks of SS as well as the 2020 CML guidelines from the ELN (16), now recommending NGS as the standard diagnostic test for *ABL1* KD analysis, prompted the unit to look for an assay that is both affordable and requires minimal hands-on time for scientist.

#### **4.1 Selection of iSeq™ 100 Sequencing System:**

The basis for choosing the iSeq™ 100 Sequencing System is the multi-fold. The analyser itself is small, requiring minimal benchtop space which is an important consideration when selecting an analyser for a new test. Secondly the cost of the analyser as well as the cost of performing the test. Setting up a new sequencing technology for a laboratory is a costly exercise. Placement of the analyser as well as reagent costs, scientist time and consumables should be considered. The unit performed a cost analysis, comparing SS to NGS (see Table 1.2). As can be seen below, the cost of performing NGS for *ABL1* KD mutational analysis, is estimated at 25% less than performing SS.

In our setting where resources are already limited, cost is a very important consideration of setting up any test. Thus NGS is shown below to be the more cost effective tool in both monetary value and scientist hands-on time.

#### **4.2 Evaluation of PCR protocols and sequencing of a retrospective cohort:**

As described in the method section, a number of PCR reactions were performed in order to optimise generation of the *ABL1* fragment. The initial failed sequencing run was a minor set-back in evaluating *ABL1* KD analysis on the iSeq™ 100 Sequencing System. This failure allowed the team to assess the cause of the failure and where methods could be improved to obtain a successful result. The first PCR amplification product using primers the unit used routinely for SS, yielded a PCR product with non-specific additional bands and off-target gene amplification, leading to an initial failed sequencing run. Higher specificity of the target of interest was needed to prevent amplification sequences that are not wanted for sequencing. In short,

although PCR primers are designed to target specific genes, a degree of overlap may occur as sequence homology is present across the entire genome (60). This can lead to amplification of “off-target” genomic sequences recorded during analysis of the sequencing run. The degree of non-specific amplification is illustrated in Figure 3.1 where numerous spurious bands can be seen in each sample.

Changing the approach to a different set of primers and mostly using the semi nested *BCR-ABL1* product as target for sequencing yielded successful results with NGS (50), concurring with those found on initial SS of the patients’ samples.

#### **4.3 Results of NGS and comparison to SS:**

Sequencing of the *ABL1* fragment during the second run of NGS showed good results with 23 out of 24 samples detecting mutations that concur with those noted to present with SS. In addition to the good concordance obtained, the average depth of coverage is noted to be ~5000.

Sample 1071/17 was the only sample that showed a very poor depth of coverage of <130. Despite the low depth the mutation was still noted to be present when the raw data was analysed. In a routine diagnostic environment this result would not be reported but rerun to try and obtain a better result. For samples 3038/19 and 3475/15, mutations were not seen during analysis of the VCF file, however on reanalysis of the BAM file both mutations were noted to be present. SS files of these samples were also reviewed to assess whether these mutations were present.

The only sample that did not show concordance with SS was sample 2148/19. Sample 2148/19 showed a deletion to be present in the *ABL1* fragment of this patient on analysis with SS. Deletions causing TKI resistance are exceedingly rare, and the function of this specific deletion is not described in the literature. Owing to retrospective analysis of the sample cohort this finding may be explained by insufficient RNA quality and/or quantity after initial testing was performed. In addition, the possibility that this mutation may not have truly been present should be considered as error introduction during sample preparation may occur. To confirm

whether this mutation is present, one can sequence this sample on an already validated analyser and compare the results. Unfortunately, no additional sample was left for this case and the additional cost of performing further analysis also precludes further information.

For some samples additional mutations were detected with NGS at low levels. Currently the accuracy of these mutations cannot be validated as these samples would have to be sequenced on an already validated instrument and be confirmed to be true. In addition, a large prospective study performed by Soverini *et al.* showed that false-positive and false-negative results may occur between 1-3% and thus their recommendation for VAF is 3% (61).

#### **4.4 Further investigations needed to the implement NGS in the routine diagnostic laboratory.**

This study proved that the iSeq™ 100 Sequencing System can be used for the detection of *ABL1* KD mutations. However, owing to the small sample size and the retrospective nature of the samples, it will be imperative to assess the sensitivity, and specificity, of this assay with real time samples referred to our laboratory in the near future.

Samples from our cohort were samples that had already been tested with SS and then frozen in a -80°C freezer as cDNA can be stored in these conditions for several years. Despite these conditions, we cannot accurately estimate if any degradation had taken place or whether prior to storage any contamination of the samples may have occurred. In addition, as freezers fill up with patient samples, boxes are moved to other freezers to make space for newer samples, this change in temperature may further influence sample integrity.

In order to validate this analyser and method going forward, a prospective study should be performed where patient samples requested for sequencing are subjected to both SS and NGS simultaneously. This allows samples to be processed according to the same method, improves the integrity of the samples and eliminate variable

parameter across the two methods. The validity of testing prospective samples has already been shown in the NEXT-in-CML study by Soverini *et al.* This is a prospective multicentre study, assessing the frequency and clinical relevance of low-level *ABL1* mutations. In addition, this team of collaborators evaluated the cost and TAT of NGS in the routine diagnostic setting in comparison to SS. A series of 236 CML patients on TKI's were tested. All these patients either had a "failure" or "warning" response. This study showed the superiority of NGS in those 51 patients who tested negative for mutations by SS, showed low-level mutations that were detected by NGS. In addition, in 60 patients where SS had detected mutations, additional low-level mutations were detected by NGS. These low-level mutations were clinically relevant in that their presence influenced TKI choice (61).

Implementation of this method is important because *ABL1* KD domain mutations have been shown to be the most important cause of resistance to TKI therapy and it is recommended to perform mutational analysis before changing therapy (16).

Multiple drug options are available for patients who don't show adequate response to their current regimen, however certain mutations are more sensitive to specific drugs than others (12, 32).

#### **4.5 Limitations:**

##### **4.5.1 Potential for introduction of error:**

A number of PCR steps are undertaken during preparation of the sample and library preparation of sequencing. During the amplification process the potential to introduce error into the sample is always present with the function of the polymerase. If an incorrect base is inserted during the PCR reactions, this may lead to a false positive or false negative result.

As mentioned above, retrospective samples were used in this study. The cDNA quality and quantity for each sample may have been affected by storage and thawing procedure during processing.

#### 4.5.2 Small sample size:

Due to the cost and limitation of only 24 samples on the kit, the sample size of this evaluation would not yield statistically significant results. This study is however the pilot study for ongoing research into utilising NGS for *ABL1* sequencing. With ongoing planned runs, enough samples will be run for statistically significant results.

### 5 CONCLUSION:

---

The Illumina® iSeq™ 100 Sequencing System for detection of *ABL1* KD mutations, using the semi nested *BCR-ABL1* PCR product was successful in detecting 23/24 mutations previously detected by SS. Currently this study does not allow for a full validation owing to the retrospective nature of samples and a prospective assessment is planned to validate and implement NGS for this purpose.

Multiple other international studies have shown NGS to be a reliable, accurate and sensitive diagnostic tool for this purpose. With both the European LeukaemiaNet and the National Comprehensive Cancer Networks now recommending NGS in CML for *ABL1* KD mutational analysis, it is imperative that the SCGU laboratory, a large centre which services multiple clinical haematology oncology centres across South Africa, keep up to date with current recommendations. NGS offers accurate and early detection of clinically relevant mutations, much earlier and at much lower levels than the previous gold standard, SS. Not only is NGS less labour intensive but direct cost analysis comparing SS with NGS showed a decreased cost per sample for NGS. Implementation of this test will increase the turnaround time, decrease scientist hands-on time, and reduce the cost per test.

In addition, once TKI suboptimal response or failure is detected, earlier detection of sensitive small clones may allow clinicians to switch to a more appropriate line of therapy, thereby inducing patients to a longer and more durable remission.

## 6 REFERENCES

---

1. Swerdlow SH CE, Harris NL, Jaffe SE, Pileri SA, Stein H, Thiele J, (Eds):. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (Revised 4th Edition). Revised 4th Edition ed: IARC: Lyon 2017; 2017.
2. NICD. Cancer in South Africa: NICD; 2019 [Available from: [https://www.nicd.ac.za/wp-content/uploads/2021/12/NCR\\_Path\\_2019\\_Full\\_Report\\_8dec2021.pdf](https://www.nicd.ac.za/wp-content/uploads/2021/12/NCR_Path_2019_Full_Report_8dec2021.pdf).
3. Rowley JD. Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature*. 1973;243(5405):290-3.
4. Kurzrock R, Gutterman JU, Talpaz M. The Molecular Genetics of Philadelphia Chromosome-Positive Leukemias. *N Engl J Med*. 1988;319(15):990-8.
5. Jabbour E, Kantarjian H. Chronic myeloid leukemia: 2018 update on diagnosis, therapy and monitoring. *Am J Hematol*. 2018;93(3):442-59.
6. Quintas-Cardama A, Cortes J. Molecular biology of bcr-abl1-positive chronic myeloid leukemia. *Blood*. 2009;113(8):1619-30.
7. Melo JV. 2 BCR-ABL gene variants. *Baillière's Clinical Haematology*. 1997;10(2):203-22.
8. Advani AS, Pendergast AM. Bcr-Abl variants: biological and clinical aspects. *Leukemia Research*. 2002;26(8):713-20.
9. Yeung DT, Hughes TP. Chronic Myeloid Leukaemia. *Postgraduate Haematology* 2015. p. 419-37.
10. Silver RT, Woolf SH, Hehlmann R, Appelbaum FR, Anderson J, Bennett C, et al. An Evidence-Based Analysis of the Effect of Busulfan, Hydroxyurea, Interferon, and Allogeneic Bone Marrow Transplantation in Treating the Chronic Phase of Chronic Myeloid Leukemia: Developed for the American Society of Hematology: Presented in part at the Education Session of the American Society of Hematology, December 5, 1998, Miami Beach, FL. *Blood*. 1999;94(5):1517-36.
11. Druker BJ, Lydon NB. Lessons learned from the development of an Abl tyrosine kinase inhibitor for chronic myelogenous leukemia. *J Clin Invest*. 2000;105(1):3-7.
12. Soverini S, Hochhaus A, Nicolini FE, Gruber F, Lange T, Saglio G, 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(5):1208-15.
13. Tokarski JS, Newitt JA, Chang CY, Cheng JD, Wittekind M, Kiefer SE, et al. The structure of Dasatinib (BMS-354825) bound to activated ABL kinase domain elucidates its inhibitory activity against imatinib-resistant ABL mutants. *Cancer Res*. 2006;66(11):5790-7.
14. Cortes JE, Saglio G, Kantarjian HM, Baccarani M, Mayer J, Boqué C, et al. Final 5-Year Study Results of DASISION: The Dasatinib Versus Imatinib Study in Treatment-Naïve Chronic Myeloid Leukemia Patients Trial. *J Clin Oncol*. 2016;34(20):2333-40.
15. Baccarani M, Castagnetti F, Gugliotta G, Palandri F, Soverini S, European Leukemia N. Response definitions and European Leukemianet Management recommendations. *Best Pract Res Clin Haematol*. 2009;22(3):331-41.
16. Hochhaus A, Baccarani M, Silver RT, Schiffer C, Apperley JF, Cervantes F, et al. European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia. *Leukemia*. 2020;34(4):966-84.
17. Soverini S, De Benedittis C, Mancini M, Martinelli G. Best Practices in Chronic Myeloid Leukemia Monitoring and Management. *Oncologist*. 2016;21(5):626-33.

18. Cross NC, White HE, Colomer D, Ehrencrona H, Foroni L, Gottardi E, et al. Laboratory recommendations for scoring deep molecular responses following treatment for chronic myeloid leukemia. *Leukemia*. 2015;29(5):999-1003.
19. Cross NC, White HE, Muller MC, Saglio G, Hochhaus A. Standardized definitions of molecular response in chronic myeloid leukemia. *Leukemia*. 2012;26(10):2172-5.
20. Hughes T, Branford S. Molecular monitoring of BCR-ABL as a guide to clinical management in chronic myeloid leukaemia. *Blood Rev*. 2006;20(1):29-41.
21. Hughes T, Deininger M, Hochhaus A, Branford S, Radich J, Kaeda J, et al. Monitoring CML patients responding to treatment with tyrosine kinase inhibitors: review and recommendations for harmonizing current methodology for detecting BCR-ABL transcripts and kinase domain mutations and for expressing results. *Blood*. 2006;108(1):28-37.
22. Yu S, Cui M, He X, Jing R, Wang H. A review of the challenge in measuring and standardizing BCR-ABL1. *Clinical Chemistry and Laboratory Medicine (CCLM)*. 2017;55(10):1465-73.
23. Deininger MW, Shah NP, Altman JK, Berman E, Bhatia R, Bhatnagar B, et al. Chronic Myeloid Leukemia, Version 2.2021, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw*. 2020;18(10):1385-415.
24. Mahon F-X. Discontinuation of TKI therapy and 'functional' cure for CML. *Best Practice & Research Clinical Haematology*. 2016;29(3):308-13.
25. Efficace F, Baccarani M, Breccia M, Alimena G, Rosti G, Cottone F, et al. Health-related quality of life in chronic myeloid leukemia patients receiving long-term therapy with imatinib compared with the general population. *Blood*. 2011;118(17):4554-60.
26. Williams LA, Garcia Gonzalez AG, Ault P, Mendoza TR, Sailors ML, Williams JL, et al. Measuring the symptom burden associated with the treatment of chronic myeloid leukemia. *Blood*. 2013;122(5):641-7.
27. Alikian M, Gerrard G, Subramanian PG, Mudge K, Foskett P, Khorashad JS, et al. BCR-ABL1 kinase domain mutations: methodology and clinical evaluation. *Am J Hematol*. 2012;87(3):298-304.
28. 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(27):5426-35.
29. Cortes J, Jabbour E, Kantarjian H, Yin CC, Shan J, O'Brien S, et al. Dynamics of BCR-ABL kinase domain mutations in chronic myeloid leukemia after sequential treatment with multiple tyrosine kinase inhibitors. *Blood*. 2007;110(12):4005-11.
30. Schindler T, Bornmann W, Pellicena P, Miller WT, Clarkson B, Kuriyan J. Structural Mechanism for STI-571 Inhibition of Abelson Tyrosine Kinase. *Science*. 2000;289(5486):1938-42.
31. Eide CA, O'Hare T. Chronic myeloid leukemia: advances in understanding disease biology and mechanisms of resistance to tyrosine kinase inhibitors. *Curr Hematol Malig Rep*. 2015;10(2):158-66.
32. Cortes JE, Kantarjian H, Shah NP, Bixby D, Mauro MJ, Flinn I, et al. Ponatinib in refractory Philadelphia chromosome-positive leukemias. *N Engl J Med*. 2012;367(22):2075-88.
33. Kuwazuru Y, Yoshimura A, Hanada S, Ichikawa M, Saito T, Uozumi K, et al. Expression of the multidrug transporter, P-glycoprotein, in chronic myelogenous leukaemia cells in blast crisis. *British Journal of Haematology*. 1990;74(1):24-9.
34. Galimberti S, Cervetti G, Guerrini F, Testi R, Pacini S, Fazzi R, et al. Quantitative molecular monitoring of BCR-ABL and MDR1 transcripts in patients with chronic myeloid leukemia during Imatinib treatment. *Cancer Genet Cytogenet*. 2005;162(1):57-62.

35. Thomas J, Wang L, Clark RE, Pirmohamed M. Active transport of imatinib into and out of cells: implications for drug resistance. *Blood*. 2004;104(12):3739-45.
36. Crossman LC, Druker BJ, Deininger MWN. hOCT 1 and resistance to imatinib. *Blood*. 2005;106(3):1133-4.
37. Sakata T, Anzai N, Shin HJ, Noshiro R, Hirata T, Yokoyama H, et al. Novel single nucleotide polymorphisms of organic cation transporter 1 (SLC22A1) affecting transport functions. *Biochem Biophys Res Commun*. 2004;313(3):789-93.
38. O'Dwyer ME, Mauro MJ, Blasdel C, Farnsworth M, Kurilik G, Hsieh Y-C, et al. Clonal evolution and lack of cytogenetic response are adverse prognostic factors for hematologic relapse of chronic phase CML patients treated with imatinib mesylate. *Blood*. 2004;103(2):451-5.
39. Cortes JE, Talpaz M, Giles F, O'Brien S, Rios MB, Shan J, et al. Prognostic significance of cytogenetic clonal evolution in patients with chronic myelogenous leukemia on imatinib mesylate therapy. *Blood*. 2003;101(10):3794-800.
40. Copland M, Hamilton A, Elrick LJ, Baird JW, Allan EK, Jordanides N, et al. Dasatinib (BMS-354825) targets an earlier progenitor population than imatinib in primary CML but does not eliminate the quiescent fraction. *Blood*. 2006;107(11):4532-9.
41. Shih AH, Abdel-Wahab O, Patel JP, Levine RL. The role of mutations in epigenetic regulators in myeloid malignancies. *Nat Rev Cancer*. 2012;12(9):599-612.
42. Roche-Lestienne C, Marceau A, Labis E, Nibourel O, Coiteux V, Guilhot J, et al. Mutation analysis of TET2, IDH1, IDH2 and ASXL1 in chronic myeloid leukemia. *Leukemia*. 2011;25(10):1661-4.
43. Soverini S, Abruzzese E, Bocchia M, Bonifacio M, Galimberti S, Gozzini A, et al. Next-generation sequencing for BCR-ABL1 kinase domain mutation testing in patients with chronic myeloid leukemia: a position paper. *J Hematol Oncol*. 2019;12(1):131.
44. Schmitt MW, Pritchard JR, Leighow SM, Aminov BI, Beppu L, Kim DS, et al. Single-Molecule Sequencing Reveals Patterns of Preexisting Drug Resistance That Suggest Treatment Strategies in Philadelphia-Positive Leukemias. *Clinical Cancer Research*. 2018;24(21):5321-34.
45. Hughes TP, Hochhaus A, Branford S, Muller MC, Kaeda JS, Foroni L, et al. Long-term prognostic significance of early molecular response to imatinib in newly diagnosed chronic myeloid leukemia: an analysis from the International Randomized Study of Interferon and STI571 (IRIS). *Blood*. 2010;116(19):3758-65.
46. Soverini S, De Benedittis C, Polakova KM, Linhartova J, Castagnetti F, Gugliotta G, et al. Next-generation sequencing for sensitive detection of BCR-ABL1 mutations relevant to tyrosine kinase inhibitor choice in imatinib-resistant patients. *Oncotarget*. 2016;7(16):21982-90.
47. Machova Polakova K, Kulvait V, Benesova A, Linhartova J, Klamova H, Jaruskova M, et al. Next-generation deep sequencing improves detection of BCR-ABL1 kinase domain mutations emerging under tyrosine kinase inhibitor treatment of chronic myeloid leukemia patients in chronic phase. *J Cancer Res Clin Oncol*. 2015;141(5):887-99.
48. Kizilers A, Crisà E, Lea N, Passera R, Mian S, Anwar J, et al. Effect of low-level BCR-ABL1 kinase domain mutations identified by next-generation sequencing in patients with chronic myeloid leukaemia: a population-based study. *The Lancet Haematology*. 2019;6(5):e276-e84.
49. Petrackova A, Vasinek M, Sedlarikova L, Dyskova T, Schneiderova P, Novosad T, et al. Standardization of Sequencing Coverage Depth in NGS: Recommendation for Detection of Clonal and Subclonal Mutations in Cancer Diagnostics. *Front Oncol*. 2019;9:851.

50. Soverini S, De Benedittis C, Machova Polakova K, Brouckova A, Horner D, Iacono M, et al. Unraveling the complexity of tyrosine kinase inhibitor-resistant populations by ultra-deep sequencing of the BCR-ABL kinase domain. *Blood*. 2013;122(9):1634-48.
51. Izzo B, Gottardi EM, Errichiello S, Daraio F, Barate C, Galimberti S. Monitoring Chronic Myeloid Leukemia: How Molecular Tools May Drive Therapeutic Approaches. *Front Oncol*. 2019;9:833.
52. Linnarsson S. Recent advances in DNA sequencing methods – general principles of sample preparation. *Experimental Cell Research*. 2010;316(8):1339-43.
53. Kia A, Gloeckner C, Osothprarop T, Gormley N, Bomati E, Stephenson M, et al. Improved genome sequencing using an engineered transposase. *BMC Biotechnology*. 2017;17(1):6.
54. Cross NCP. Standardisation of molecular monitoring for chronic myeloid leukaemia. *Best Practice & Research Clinical Haematology*. 2009;22(3):355-65.
55. van Dongen JJM, Macintyre EA, Gabert JA, Delabesse E, Rossi V, Saglio G, et al. Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease. *Leukemia*. 1999;13(12):1901-28.
56. Branford S, Cross NCP, Hochhaus A, Radich J, Saglio G, Kaeda J, et al. Rationale for the recommendations for harmonizing current methodology for detecting BCR-ABL transcripts in patients with chronic myeloid leukaemia. *Leukemia*. 2006;20(11):1925-30.
57. Hochhaus A, La Rosée P. Imatinib therapy in chronic myelogenous leukemia: strategies to avoid and overcome resistance. *Leukemia*. 2004;18(8):1321-31.
58. Bishop MJ. Guide to human genome computing / edited by Martin J. Bishop. 2nd ed. ed. San Diego, Calif: Academic Press; 1998.
59. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114-20.
60. Yee J-K. Off-target effects of engineered nucleases. *The FEBS Journal*. 2016;283(17):3239-48.
61. Soverini S, Bavaro L, De Benedittis C, Martelli M, Iurlo A, Orofino N, et al. Prospective assessment of NGS-detectable mutations in CML patients with nonoptimal response: the NEXT-in-CML study. *Blood*. 2020;135(8):534-41.
62. Illumina. Nextera XT DNA Library Prep 2019 [Available from: [https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry\\_documentation/samplepreps\\_nextera/nextera-xt/nextera-xt-library-prep-reference-guide-15031942-05.pdf](https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/samplepreps_nextera/nextera-xt/nextera-xt-library-prep-reference-guide-15031942-05.pdf)].

## 7 Appendices:

### 7.1 Appendix A: Reagents used for PCR reactions.

Table 7.1 PCR reagents used to generate cDNA, first round *BCR-ABL1* and second round *ABL1*.

| PCR reagents                                    | cDNA                                                                                                                                                                                                                                                                     |
|-------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ezDNase buffer                                  | 1.25 uL                                                                                                                                                                                                                                                                  |
| ezDNase enzyme                                  | 1.25 uL                                                                                                                                                                                                                                                                  |
| SuperScript Master Mix                          | 5 uL                                                                                                                                                                                                                                                                     |
| Molecular Grade water                           | 7.5 uL                                                                                                                                                                                                                                                                   |
| RNA                                             | 1 ug                                                                                                                                                                                                                                                                     |
| <b>Thermocycling conditions</b>                 | 25°C for 10 minutes, 42°C for 45 minutes, 99°C fir 3 min and 4°C hold                                                                                                                                                                                                    |
| <b>PCR reagents first round <i>BCR-ABL1</i></b> | <b><i>BCR-ABL1</i></b>                                                                                                                                                                                                                                                   |
| AmpliTaq Gold 360 Master Mix                    | 12.5 uL                                                                                                                                                                                                                                                                  |
| CML F2 (RT) 10 pmol/uL primer                   | 2 uL                                                                                                                                                                                                                                                                     |
| BCR/ <i>ABL1</i> seq R1 or R3 10 pmol/uL primer | 2 uL                                                                                                                                                                                                                                                                     |
| Molecular grade water                           | 6.5 uL                                                                                                                                                                                                                                                                   |
| cDNA                                            | 4uL                                                                                                                                                                                                                                                                      |
| <b>Thermocycling conditions</b>                 | An initial denaturation step was carried out at 95°C for 7min, followed by 35 cycles (denaturation at 95°C for 30 seconds, annealing at 62°C for 30 seconds and extension at 72°C for 1.5 minutes) with final extension at 72°C for 7 minutes and 4°C hold indefinitely. |
| <b>PCR reagents second round <i>ABL1</i></b>    | <b><i>ABL1</i></b>                                                                                                                                                                                                                                                       |
| AmpliTaq Gold 360 Master Mix                    | 12.5 uL                                                                                                                                                                                                                                                                  |
| <i>ABL1</i> F2 10 pmol/uL primer                | 1 uL                                                                                                                                                                                                                                                                     |
| BCR/ <i>ABL1</i> seq R1 or R3 10 pmol/uL primer | 1 uL                                                                                                                                                                                                                                                                     |
| Molecular grade water                           | 6.5 uL                                                                                                                                                                                                                                                                   |
| First round product ( <i>BCR-ABL1</i> )         | 1 uL                                                                                                                                                                                                                                                                     |
| <b>Thermocycling conditions</b>                 | An initial denaturation step was carried out at 95°C for 7min, followed by 35 cycles (denaturation at 95°C for 30 seconds, annealing at 62°C for 30 seconds and                                                                                                          |

|  |                                                                                                          |
|--|----------------------------------------------------------------------------------------------------------|
|  | extension at 72°C for 1.5 minutes) with final extension at 72°C for 7 minutes and 4°C hold indefinitely. |
|--|----------------------------------------------------------------------------------------------------------|

## **7.2 Appendix B: Method for making an agarose gel.**

### **2% Agarose Gel**

- Mix 2g agarose with 100ml 1x Tris-Borate Ethylenediaminetetraacetic acid (TBE) and boil until agarose is completely dissolved. Allow to cool for 25 minutes. Add 10 $\mu$ l (10 000X) Gel Red to the solution. Pour into gel casting tray to set.

### **10X TBE (Tris/Borate/EDTA)**

- Dissolve 108 g of Tris base and 55 g boric acid in 800 ml of dH<sub>2</sub>O. Add 40 ml of 0.5 M EDTA (pH 8.0) and make up to 1 litre. The pH of this stock should be 8.3 1X TBE

- Mix 100 ml of 10X TBE stock and 900 ml of dH<sub>2</sub>O to make 1 litre of 1X TBE solution

Approximately 5 $\mu$ l of DNA, RNA or PCR product was mixed with 3 $\mu$ l of ficoll dye and were subsequently loaded on a 2% agarose gel with a molecular weight ladder and resolved at 90 volts for 60 minutes in a gel tank with TBE buffer

### 7.3 Appendix C: Materials used for Illumina® library preparation and sequencing.

#### EQUIPMENT, REAGENTS AND CONSUMABLES

##### Equipment

4°C (2°C to 8°C) fridge  
-20°C (-15°C to -25°C) freezer  
-70°C to -80°C freezer  
Illumina® iSeq™ 100 Sequencing System  
High-quality thermal cycler\*  
Calibrated pipettes  
Qubit fluorometer 2.0 or later  
Minispin centrifuge  
NHLS computer workstation with access to local NHLS intranet

##### Reagents and Consumables

| <u>Reagent or Consumable</u>                         | <u>Storage</u>    |
|------------------------------------------------------|-------------------|
| <u>conditions</u>                                    |                   |
| Virkon                                               | RT (20°C to 30°C) |
| 70% Isopropanol                                      | RT (20°C to 30°C) |
| DNA Away                                             | RT (20°C to 30°C) |
| RNase Away                                           | RT (20°C to 30°C) |
| Molecular grade water                                | RT (20°C to 30°C) |
| Pipette tips                                         | RT (20°C to 30°C) |
| Microcentrifuge tubes (1.5 to 2.0 ml)                | RT (20°C to 30°C) |
| Adhesive PCR plate seals OR ABI Optical 8-cap strips | RT (20°C to 30°C) |
| Gloves                                               | RT (20°C to 30°C) |
| Paper towel                                          | RT (20°C to 30°C) |
| Nextera XT Library Prep Kit (24 samples)             |                   |
| Box 1/2                                              |                   |
| Amplicon Tagment Mix                                 | -25°C to -15°C    |

|                                        |                        |
|----------------------------------------|------------------------|
| Tagment DNA Buffer                     | -25°C to -15°C         |
| Hybridization Buffer                   | -25°C to -15°C         |
| Nextera PCR Master Mix                 | -25°C to -15°C         |
| Resuspension Buffer                    | -25°C to -15°C         |
| Library Normalization Additives 1      | -25°C to -15°C         |
| Library Normalization Wash 1           | 2°C to 8°C             |
| Box2/2                                 |                        |
| Neutralize Tagment Buffer              | Room temperature       |
| Library Normalization Beads            | 2°C to 8°C             |
| Library Normalization Storage Buffer 1 | Room temperature       |
| Superscript IV VILO Master Mix         |                        |
| Superscript IV VILO Master Mix         | -20°C (-15°C to -25°C) |
| Qubit RNA HS assay, consisting of:     |                        |
| Buffer and dye                         | RT (20°C to 30°C)      |
| Standards                              | 4°C (2°C to 8°C)       |
| Qubit dsDNA HS assay, consisting of:   |                        |
| Buffer and dye                         | RT (20°C to 30°C)      |
| Standards                              | 4°C (2°C to 8°C)       |
| Qubit Assay tubes                      | RT (20°C to 30°C)      |

## **7.4 Appendix D: Manufacturers' protocol (62).**

### **Tagmenting of genomic DNA:**

Nextera XT DNA Library Preparation kit makes use of transposome technology to fragment and at the same time tag DNA with special adapter sequences. A 96-well PCR plate was used for sample preparation. Samples volume was calculated to contain 10ng of cDNA and the difference in volume was equalised by adding nuclease-free water to bring the total volume to 30uL. A tagmentation master mix was made by mixing a total of 264uL (11uL/sample) of bead-linked transposomes with 264uL (11uL/sample) tagmentation buffer 1 and vortexing thoroughly to resuspend. 20uL Tagmentation master mix was pipetted into each well, followed by 10x pipetting to allow for resuspension.

The sealed plate was then placed on the thermal cycler and run on the following conditions: the preheat option is chosen and the temperature set to 100°C. Annealing is set at 55°C for 15 min with hold indefinitely at 10°C.

### **Post Tagmentation Cleanup:**

After tagmentation, 10uL of tagment stop buffer is added to the tagmentation reaction and slowly pipetted, 10 times to allow for resuspension of the beads. The plate is once again placed on the thermal cycler and run with on the following conditions: 37°C for 15 minutes and hold indefinitely at 10°C. The plate is then placed on a magnetic stand and allowed to sit for ~3 minutes, until the liquid is clear. Once the liquid is clear, the supernatant is removed from each sample with a pipette and discarded. A washing step is then applied to each sample as follows: 100uL of tagment wash buffer is added directly to the beads in each sample using a deliberately slow pipetting technique to avoid foaming of the tagment wash buffer, and allow for full resuspension. After resuspension the plate is once again placed on the magnetic stand and allowed to sit for ~3 minutes until the liquid is clear. The supernatant is once again discarded and the entire process repeated for a second time. The plate is left on the magnetic stand, until needed in the amplifying of tagmented DNA step.

### **Amplification of tagmented DNA:**

In this step the tagmented DNA is amplified and patient specific indexes (i7 and i5) are added. These indexes are required for cluster generation. A PCR master mix is prepared by combining the following reagents for the amount of samples for analysis: 22ul of enhanced PCR mix per patient sample with 22uL nuclease free water, per patient sample. The master mix is then centrifuged for 10 seconds at 280 x g. The supernatant left in the plate on the magnetic stand is removed. After removing the plate from the magnetic stand, 40uL of PCR master mix is added onto the beads in each sample well and pipetted until the beads are fully resuspended. The plate is sealed and centrifuged at 280 x g for 3 seconds.

5uL of the i7 and the i5 adapters are added to each sample and mixed, then centrifuged at 280 x g for 30 seconds and run on the thermal cycler under the following conditions: the lid is preheated at 100°C, then 68°C for 3 minutes and 98°C for 3 minutes. 8 cycles of 98°C for 45 seconds, 62°C for 30 seconds and 68°C for 2 minutes are run, then 68°C for 1 minute and hold indefinitely at 10°C.

### **Library clean up:**

For the library clean up double-sided bead purification is used to purify the amplified libraries. To start, the plate is centrifuged at 280 x g for 1 minute to collect the contents of each sample at the bottom of the well. The plate is rested on the magnetic stand for ~5 minutes. 45uL supernatant from each well is transferred to the corresponding well of a new PCR plate, which is then vortexed. 40uL of nuclease water is then added to each well as well as 45uL sample purification beads and resuspended. The plate is sealed and incubated at room temperature for 5 minutes and transferred onto the magnetic stand for 5 minutes. Stock, undiluted sample purification beads are resuspended and 15ul added to a new PCR plate. 125uL of supernatant from well of the first plate is pipetted into the corresponding well of the second plate, which contains the 15uL of sample purification beads. Each well is pipetted 10 times to allow for adequate mixing and the first plate is discarded. The PCR plate is incubated at room temperature for 5 minutes, then transferred to the magnetic stand and left for another 5 minutes.

The supernatant from each well is removed and so that a washing step can commence. The washing step is as follows: with the plate still on the magnetic stand, add 200uL of fresh 80% ethanol without mixing. Allow the samples to stand for 30 seconds, then without disturbing the beads remove and discard the supernatant. After the washing step, a 20uL pipette is used to remove and discard the residual ethanol, and the plate is left to air-dry on the magnetic stand for 5 minutes. Beads are resuspended with 32uL resuspension buffer and incubated at room temperature for 2 minutes, then placed on the magnetic stand for 2 minutes until the liquid is clear. 30uL of the supernatant from each well is once again transferred to the corresponding well of a new PCR plate.

### **Library pooling:**

Each sample was quantified before pooling the libraries using the Qubit fluorometer, using the method described above. The 24 samples that were used for library preparations were quantified, only 20 of these samples were eligible for sequencing due to 4 samples not having a high enough cDNA concentration for sequencing. Libraries have to be diluted to the starting concentration for the Illumina® iSeq™ 100 Sequencing System after which the libraries are ready to be denatured and diluted to the final loading concentration onto the sequencing cartridge.

The molarity of the value of the libraries are calculated using the following formula:

$$\frac{\text{Concentration in ng/uL} \times 10^6}{660\text{g/mol} \times \text{average size of library (500bp)}} = [\text{nM}]$$

Using the molarity value, the volume of resuspension buffer and library needed to dilute to the starting concentration is calculated. Libraries were diluted 1nM which equates to 100pM final loading concentration of each sample. 10uL of each diluted library is then added to a tube to create a multiplexed library pool, ready for loading onto the sequencing cartridge. The pooled libraries are then loaded onto the sequencing cartridge by pipetting 20uL of the pooled library to the bottom of the

reservoir of the cartridge and loading the cartridge and the flow cell into the iSeq100 sequencing system.

## 7.5 Appendix E: Snapshot of sequences obtained from the failed sequencing run.

| A1    | A       | B  | C        | D        | E    | F      | G                                                                        | H                                 | I                                                     | J  |
|-------|---------|----|----------|----------|------|--------|--------------------------------------------------------------------------|-----------------------------------|-------------------------------------------------------|----|
| CHROM | POS     | ID | REF      | ALT      | QUAL | FILTER | INFO                                                                     | FORMAT                            |                                                       |    |
| chr1  | 1050466 |    | G        | A        | .    | .      | DP=1MQ=250.00SQ=7.95;GermlineQuality=6.98;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.95:0.1:1.000:0.1:0.0:1.0:0.0:1                  | 43 |
| chr1  | 1086657 |    | C        | T        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.1:0.0:1.0:0.0:1                  |    |
| chr1  | 1295954 |    | A        | G        | .    | .      | DP=1MQ=250.00SQ=7.94;GermlineQuality=6.98;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.94:0.1:1.000:0.1:0.0:1.0:0.0:1                  |    |
| chr1  | 1297051 |    | CCA      | C        | .    | .      | DP=1MQ=250.00SQ=4.35;GermlineQuality=7.87;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:4.35:0.1:1.000:0.1:0.0:1.0:0.0:1                  |    |
| chr1  | 1532105 |    | T        | C        | .    | .      | DP=1MQ=208.00SQ=7.94;GermlineQuality=6.98;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.94:0.1:1.000:0.1:0.0:1.0:0.0:1                  |    |
| chr1  | 1517354 |    | T        | C        | .    | .      | DP=1MQ=124.00SQ=7.98;GermlineQuality=7.00;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.98:0.1:1.000:0.0:1.0:0.0:1.0:0.1                |    |
| chr1  | 1665581 |    | C        | G        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.0:1.0:0.0:1.0:0.1                |    |
| chr1  | 1752206 |    | C        | T        | .    | .      | DP=2MQ=213.38SQ=5.69;GermlineQuality=5.34;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:5.69:1.10:500:1.1:0.0:2.1:0.0:1.0:0.1             |    |
| chr1  | 1778660 |    | C        | T        | .    | .      | DP=1MQ=250.00SQ=0.00;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB:PS | 0/1:0.00:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1:0.1778660  |    |
| chr1  | 1778667 |    | A        | G        | .    | .      | DP=1MQ=250.00SQ=0.00;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB:PS | 0/1:0.00:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1:0.1778660  |    |
| chr1  | 1778670 |    | G        | A        | .    | .      | DP=1MQ=250.00SQ=0.00;GermlineQuality=6.94;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB:PS | 0/1:0.00:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1:0.1778660  |    |
| chr1  | 2536095 |    | C        | T        | .    | .      | DP=1MQ=250.00SQ=7.98;GermlineQuality=7.00;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.98:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 2528735 |    | C        | T        | .    | .      | DP=1MQ=250.00SQ=7.82;GermlineQuality=6.88;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.82:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 3076608 |    | C        | T        | .    | .      | DP=1MQ=250.00SQ=7.99;GermlineQuality=7.01;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.99:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 3134374 |    | T        | C        | .    | .      | DP=1MQ=212.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 3147001 |    | T        | C        | .    | .      | DP=1MQ=248.00SQ=0.92;GermlineQuality=6.96;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB:PS | 0/1:0.92:0.1:1.000:0.0:0.1:1.0:0.1:0.0:0.1:0.13147001 |    |
| chr1  | 3147014 |    | A        | G        | .    | .      | DP=1MQ=211.00SQ=0.84;GermlineQuality=6.90;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB:PS | 0/1:0.84:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1:0.13147001 |    |
| chr1  | 3147074 |    | GTGTA    | G        | .    | .      | DP=1MQ=211.00SQ=6.48;GermlineQuality=5.82;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:6.48:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 3222964 |    | C        | T        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 3270678 |    | A        | AGAGGAGG | .    | .      | DP=1MQ=131.00SQ=9.65;GermlineQuality=6.09;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:6.05:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 3270752 |    | A        | G        | .    | .      | DP=1MQ=131.00SQ=7.98;GermlineQuality=6.85;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.98:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 3276395 |    | A        | G        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 3414432 |    | G        | A        | .    | .      | DP=1MQ=250.00SQ=7.96;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.96:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 3502093 |    | G        | A        | .    | .      | DP=1MQ=169.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 3598870 |    | T        | C        | .    | .      | DP=1MQ=250.00SQ=7.94;GermlineQuality=6.98;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.94:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 3864532 |    | C        | A        | .    | .      | DP=1MQ=250.00SQ=7.93;GermlineQuality=6.95;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.93:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 3864572 |    | C        | G        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 3936431 |    | G        | T        | .    | .      | DP=1MQ=243.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 3936483 |    | G        | C        | .    | .      | DP=1MQ=243.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 3979772 |    | T        | A        | .    | .      | DP=1MQ=41.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000  | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 5107393 |    | GA       | G        | .    | .      | DP=1MQ=58.00SQ=1.54;GermlineQuality=1.42;FractionInformativeReads=1.000  | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:1.54:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 5166692 |    | G        | A        | .    | .      | DP=1MQ=46.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000  | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 5239881 |    | C        | G        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 5247209 |    | T        | TGTGTGGG | .    | .      | DP=2MQ=0.00SQ=0.98;GermlineQuality=8.17;FractionInformativeReads=0.000   | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:5.99:0.0:0.1:0.000:0.0:0.0:0.0:0.0:0.0            |    |
| chr1  | 5247117 |    | AT       | A        | .    | .      | DP=1MQ=250.00SQ=1.56;GermlineQuality=1.43;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:1.56:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 5263340 |    | C        | A        | .    | .      | DP=1MQ=250.00SQ=7.98;GermlineQuality=6.98;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.98:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 5273280 |    | T        | C        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 5273324 |    | C        | T        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 5283865 |    | T        | C        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=7.00;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 5340873 |    | A        | G        | .    | .      | DP=1MQ=250.00SQ=7.96;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.96:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 5376422 |    | C        | T        | .    | .      | DP=1MQ=250.00SQ=0.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB:PS | 0/1:0.97:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1:0.5376422  |    |
| chr1  | 5376427 |    | T        | C        | .    | .      | DP=1MQ=250.00SQ=0.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB:PS | 0/1:0.97:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1:0.5376422  |    |
| chr1  | 5631644 |    | T        | G        | .    | .      | DP=1MQ=223.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 5639574 |    | G        | A        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 5635033 |    | A        | G        | .    | .      | DP=1MQ=250.00SQ=0.13;GermlineQuality=1.18;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:1.25:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 6028213 |    | T        | G        | .    | .      | DP=1MQ=250.00SQ=7.98;GermlineQuality=7.00;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.98:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 6044421 |    | G        | A        | .    | .      | DP=1MQ=250.00SQ=7.96;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.96:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 6228809 |    | C        | T        | .    | .      | DP=1MQ=142.00SQ=0.96;GermlineQuality=6.98;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB:PS | 0/1:0.96:0.1:1.000:0.0:0.1:1.0:0.1:0.0:0.1:0.6228809  |    |
| chr1  | 6228831 |    | C        | A        | .    | .      | DP=1MQ=142.00SQ=0.96;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB:PS | 0/1:0.96:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1:0.6228809  |    |
| chr1  | 6255472 |    | C        | T        | .    | .      | DP=1MQ=250.00SQ=7.96;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.96:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 6255557 |    | A        | T        | .    | .      | DP=1MQ=250.00SQ=7.96;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.96:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 6354477 |    | A        | T        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 7092574 |    | T        | C        | .    | .      | DP=1MQ=250.00SQ=7.96;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 7428293 |    | G        | C        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 7501238 |    | A        | G        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 7754631 |    | G        | C        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 8043608 |    | G        | A        | .    | .      | DP=1MQ=106.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 8132827 |    | A        | G        | .    | .      | DP=1MQ=250.00SQ=7.95;GermlineQuality=6.97;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.95:0.1:1.000:0.0:0.1:1.0:0.0:1.0:0.1            |    |
| chr1  | 8154603 |    | G        | A        | .    | .      | DP=1MQ=250.00SQ=7.97;GermlineQuality=6.99;FractionInformativeReads=1.000 | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:7.97:0.1:1.000:0.1:0.0:1.0:0.0:1.0:0.1            |    |
| chr1  | 8243248 |    | GAAGAAAG | G        | .    | .      | DP=1MQ=92.00SQ=0.73;GermlineQuality=0.68;FractionInformativeReads=1.000  | GT:SQ:AD:AF:F1R2:F2R1:DP:SB:MB    | 0/1:0.73:0.1:1.000:0.0:0.1:1.0:0.1:0.0:0.0:1          |    |

Figure 7.1 Snapshot of sequences obtained of chromosome 9.

Column A (CHROM) indicates the chromosome to which the sequence aligned; column B indicated the coordinates on the chromosome; column C indicated the unique identifying sequence (not shown in the column as this is a single sample); column D (REF) is the reference allele and indicates what was found on the reference allele; column E (ALT) is the allele found in the sample; column F (QUAL) is the quality score (not indicated); column G (FILTER) indicates whether the quality filters were passed or failed; column H (INFO) allows one to provide further information of the variants; column I (FORMAT) the GT format indicates one to expect genotypes in the following columns; column J (433) an optional individual identifier: column I indicated expected genotypes here, and is indicated in the form 0|1. 0 indicates the reference allele and 1 indicates the alternative allele.

Figure 7.1 above shows an illustration of what a VCF file from one sample looks like. In this instance sequences aligning to chromosome 1 are illustrated, however many more data entries, covering other chromosomes (not shown) are also in this file. Because sequences are aligning to all other chromosomes as well, results from this sequencing run are not adequate for verification of NGS for *ABL1* kinase domain mutational analysis.

## 7.6 Appendix F: Wits ethics clearance



R14/49 Dr T De Jager

### **HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL) CLEARANCE CERTIFICATE NO. M200567**

**NAME:** Dr T De Jager  
(Principal Investigator)

**DEPARTMENT:** School of Pathology  
Department of Molecular Medicine and Haematology  
Medical School  
University

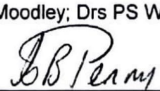
**PROJECT TITLE:** Evaluation of next generation sequencing panel for ABL1  
kinase domain mutations in patients with BCR-ABL1  
positive Chronic Myeloid Leukaemia

**DATE CONSIDERED:** 2020/05/29

**DECISION:** Approved unconditionally

**CONDITIONS:**

**SUPERVISOR:** Mr M Moodley; Drs PS Willem and N Bouwer

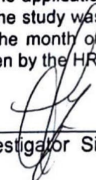
**APPROVED BY:**   
Dr CB Penny, Chairperson, HREC (Medical)

**DATE OF APPROVAL:** 2020/07/02

This clearance certificate is valid for 5 years from the date of approval. Extension may be applied for.

#### **DECLARATION OF INVESTIGATORS**

To be completed in duplicate and ONE COPY returned to the Research Office Secretary on the 3rd Floor, Phillip Tobias Building, Parktown, University of the Witwatersrand, Johannesburg.  
I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to submit details to the Committee. I agree to submit a yearly progress report. When a funder requires annual re-certification, the application date will be one year after the date when the study was initially reviewed. In this case, the study was initially reviewed in May and will therefore reports and re-certification will be due early in the month of May each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

  
Principal Investigator Signature

9/7/2020  
Date

## 7.7 Appendix G: Plagiarism/turn-it-in report cover page

2326462:MC\_V10\_turnitin.docx  
*by Tanya de Jager*

---

**Submission date:** 09-May-2022 09:36PM (UTC+0200)

**Submission ID:** 1832326942

**File name:** nments\_cae9b470-0e7b-401f-bfbe-170eac2a64e3\_MC\_V10\_turnitin.docx (19.37M)

**Word count:** 14618

**Character count:** 77062