DETECTION OF CIRCULAR RNAS IN HEPATOCELLULAR CARCINOMA TREATED WITH DIFFERENT THERAPEUTIC AGENTS USING RNA-SEQ DATA

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BURCU TAMKOÇ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
BIOTECHNOLOGY

SEPTEMBER 2021

Approval of the thesis:

DETECTION OF CIRCULAR RNAS IN HEPATOCELLULAR CARCINOMA TREATED WITH DIFFERENT THERAPEUTIC AGENTS USING RNA-SEQ DATA

submitted by **BURCU TAMKOÇ** in partial fulfillment of the requirements for the degree of **Master of Science** i**n Biotechnology, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Candan Gürakan
Head of the Department, **Biotechnology** _____

Assist. Prof. Dr. Aybar Can Acar
Supervisor, **Biotechnology, METU** _____

Assoc. Prof. Dr. Can Özen
Co-Supervisor, **Biotechnology, METU** _____

Examining Committee Members:

Prof. Dr. Tolga Can
Computer Engineering, METU _____

Assist. Prof. Dr. Aybar Can Acar
Health Informatics, METU _____

Assoc. Prof. Dr. Nurcan Tunçbağ
College of Engineering and School of Medicine, Koç University _____

Date: 08.09.2021

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name Last name : Burcu Tamkoç

Signature :

ABSTRACT

DETECTION OF CIRCULAR RNAS IN HEPATOCELLULAR CARCINOMA TREATED WITH DIFFERENT THERAPEUTIC AGENTS USING RNA-SEQ DATA

Tamkoç, Burcu
Master of Science, Biotechnology
Supervisor : Assist. Prof. Dr. Aybar Can Acar
Co-Supervisor: Assist. Prof. Dr. Can Özen

September 2021, 125 pages

Discovering therapeutic agents for hepatocellular carcinoma (HCC) pose challenges as HCC contains high heterogeneity in terms of molecular phenotypes. It is thus crucial to understand the role of circular RNAs in the treatment of HCC. Circular RNAs are small and non-coding endogenous RNAs that have single stranded and covalently bound structures. The clinical importance of circular RNAs is due to their high stability in the serum and tissue. Up to now, the detection of circular RNAs was done through specific experiments that were pre-enriched for circular isoforms. However, it is also possible to detect circular RNAs from regular RNA-Seq data using unmapped back-splice junction information. The aim of this thesis is to detect circular RNAs from the raw RNA-Seq data and to analyze their transcript level differential expression across conditions like treatment and non-treatment cell lines. The qualitative and quantitative circular RNA expression differences between the Mahlavu and HUH-7 HCC cell lines treated with sorafenib, PI3K-α, PI3K-β inhibitors and their combinations (i.e. sorafenib and PI3K-α inhibitor; and sorafenib and PI3K- β inhibitor) will be accomplished. This was be achieved by using 12 raw RNA-seq read sets. CirComPara was selected as detection and quantification method

as it allows of working simultaneously with different detection methods such as CIRI, CircExplorer2, DCC, findcirc, circRNAfinder. In total 136 and 122 reliable circular RNAs were detected in HUH-7 and Mahlavu, respectively. When HUH-7 and Mahlavu cell line were not under any treatment, 12 common circular RNAs could be detected in both cell lines. Otherwise, no common circular RNAs were detected.

Keywords: Circular RNA, Hepatocellular Carcinoma, RNA-Seq, Drug Treatment

ÖZ

FARKLI TEDAVİ AJANLARI UYGUNLANMIŞ KARACİĞER KANSERLERİNE AİT RNA DİZİLERİNDEN ÇEMBERSEL RNA'LARIN TESPİT EDİLMESİ

Tamkoç, Burcu
Yüksek Lisans, Biyoteknoloji
Tez Yöneticisi: Yrd. Doç. Dr. Aybar Can Acar
Ortak Tez Yöneticisi: Yrd. Doç. Dr. Can Özen

Eylül 2021, 125 sayfa

Karaciğer kanserinde tedavi ajanlarının keşfedilmesi bir çok zorluk içermektedir, çünkü karaciğer kanseri moleküler fenotipler açısından fazla heterojenlik içerir. Karaciğer kanseri tedavisinde çembersel RNA'ların rolünü anlamak çok önemlidir. Çembersel RNA'lar küçük, kodlanmayan endojenik, tek sarmalı, kovalent bağlı RNA yapılarıdır. Yüksek stabiliteleri sayesinde çembersel RNA'ların önemi ortaya çıkmıştır. Bu zamana kadar çembersel RNA'ların belirlenmesi çembersel RNA isoformu kullanılarak özgül deneylerle yapılıyordu. Ancak RNA-seq verisi kullanılarak ters uç birleştirme yerleri üzerinden çembersel RNA'ları belirlemek mümkündür. Bu tezin amacı bahsedilen ham ve haritanlamamış veriler ile olası ek birleşme yerleri hesaba katılarak çembersel RNA'lar bulmak ve çembersel RNA'ların farklı koşullardaki (farklı tedavi uygulanmış-tedavi uygulanmamış) transkript seviyelerini analiz etmektir. Mahlavu ve HUH-7 kanser hücre hatlarına ait 12 adet RNA-seq verisi kullanmıştır. Kullanılan RNA-seq verileri; mahlavu ve HUH-7 kanser hücre hatlarının sorafenib, PI3K-α, PI3K- β inhibitörü, ve bu ilaçların kombinasyonları olarak sorafenib ve PI3K-α inhibitörü ve sorafenib ve

PI3K-α, PI3K- β inhibitörü ile muamele edilmiş verileridir. Bir çok farklı tespit yöntemlerini; CIRI, CircExplorer2, DCC, findcirc ve circRNAfinder, aynı anda çalıştırılmasına olanak sağladığı için çembersel RNA'ların tespitinde ve kantitatif analizinde CirComPara programı kullanılmıştır. HUH-7 hücre hattında toplam 136 tane çembersel RNA tespit edilirken, Mahlavu hücre hattında ise 122 tane çembersel RNA tespit edilmiştir. İki hücre hattında herhangi bir tedavi ajan ile muamele edilmemiş iki hücre hattında ortak 12 tane çembersel RNA tespit edilmiştir; tedavi ajanlarıyla muamele edilmiş hücre hatlarında ortak bir çembersel RNA tespit edilememiştir.

Anahtar Kelimeler: Çembersel RNA, RNA-seq Verisi, Karaciğer Kanseri, Tedavi Ajanları

To my brother

ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my supervisor, Assist. Prof. Dr. Aybar Can Acar who guided me throughout this project. During the pandemic of Covid-19, his guidance helped me in studying constantly of this thesis.

Besides my advisor, I would like to offer my special thanks to Assoc. Prof. Dr. Can Özen for his encouragement and continuous support. I would also like to thank my committee members, Prof. Dr Tolga Can and Assist. Prof. Dr. Nurcan Tunçbağ for their valuable comments and suggestions.

In particular I would like to show my gratitude to my friends Berfu Safranbolulu and Ozan Çiftçi for their unconditional support during these years.

To conclude, I cannot forget to thank my family for all unconditional support in this very intense academic year. This thesis cannot be existing if they do not believe in me.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation

Single stranded circRNAs are a new class of abundant functional RNAs that have covalently closed structures, which enable them to be highly stable and found in saliva, urine and blood, as well as having neither 5'-3' polarity nor polyadenylated tail. [15, 10] In addition to stability, they have remarkable properties like abundance, evolutionary conservation, and cell/tissue specificity [10]. One of the cancer types that possibly circular RNAs (circRNAs) involve is hepatocellular carcinoma (HCC), which is the fifth most common cancer and the third leading cause of cancer-related deaths worldwide [1,2]. There are various type of risk factors, such as viral hepatitis, alcohol, hemochromatosis, obesity, and metabolic syndrome or genotoxins [3]. HCCs contain high heterogeneity in terms of molecular phenotypes. This raises major challenges in clinical management. Tumor number, tumor size, vascular invasion, and functional liver reserve are used as patient managements nowadays [4]. But patient management is not enough to diagnose properly and to select the treatment type, due to the phenotypic and genetic heterogeneity of HCCs. Although the first FDA approved drug for HCC patients at advanced stages is Sorafenib, cancer cells can resist the drug and may even increase their migratory and invasive abilities. [8] Although finding new therapeutical agents is an urgent issue for the HCC patients, understanding the resistance mechanism of the cancer cell is an important topic for the HCC as well.

They may involve the HCC related pathways like AKT/MTOR or RAS/MAPK pathways by affecting their parental gene through miRNA-mRNA interaction. Discovering the effects of circRNAs on cancer cell lines with different treatment options may give researchers greater insight.

1.2 Scope and Goal

Although there are a few studies related to investigation of the relation between HCC and circular RNAs [11, 27, 67, 68, 69, 70, 71, 72, 73, 75, 76, 77], there is limited research [8, 9] that has focused on therapeutic agents that are used in HCC in relation to circRNAs. Having a look at different aspects, we have focused on RNA-sequences of different cell lines treated with different therapeutic agents. We aimed to investigate the role of circRNAs in different cell lines that were treated with sorafenib, PI3K-α inhibitor, PI3K-β inhibitor and their combinations as sorafenib and PI3K-α, sorafenib and PI3K-β inhibition. HUH-7 and Mahlavu cell lines were selected as typical HCC and atypical cell lines, respectively. We have tried to detect common and unique circRNAs that were differentially expressed in each sample. Through circRNA-miRNA-mRNA relations, we have examined the role of gene expression regulation in two different cell lines with different treatment options.

1.3 Contribution

In this thesis, HUH-7 and Mahlavu cell lines treated with different treatment options were selected. Although RNA sequences of these cell lines were not generated for the detection of circRNAs through library preparation methods, circRNAs were detected by using total RNA sequences. CirComPara [52] was preferred as detection and quantification method since it enables researchers to use different methods in the same pipeline, in order to reduce false positives. After detection of circRNAs, just by looking at the differential expression of linear isoforms, Gene Set Enrichment Analysis has been done for each sample.

CHAPTER 2

GENERAL OVERVIEW AND HISTORICAL APPROACHES

In 1976, Circular RNA molecules were discovered in viroids which are pathogenic to many plants. Circular RNAs were later detected in eukaryote cytoplasms in 1979. In their seminal papers in 1990s, Nigro, Capel and Zophiropoulos found that circular RNAs molecules play an important role in transcripts of many genes, such as the DCC tumor suppressor gene, Sry gene and rat cytochrome P450 2C24 gene. Increasingly many different functions of circRNAs molecules, as well as a large number of new circRNAs were reported by by Salzman, Jeck, Memczak and Hansen in 2013. More recently, there has been a considerable amount of literature on circRNAs' biogenesis, potential functions, and clinical implications [10].

The report in circRNA abundancy was carried out in 2013 by Jeck et al. Specifically, they found that some exonic circRNAs are ten times more abundant than linear mRNA in vivo [66]. CircRNAs can have interaction with RNA-binding proteins, can also be translated, and play a role in gene expression regulation by acting as miRNA sponges or transcriptional regulators [10]. Although whether all circRNA can act as miRNA sponges or not is still unclear, some circRNAs have some miRNA binding sites that can reduce or increase miRNA activity via binding their specific sites [12]. Human CDR1as and ciRS-7 are the examples of circRNAs that have been experimentally validated for miRNA sponge activity [13]. In this way, they regulate gene expression in mammals' transcriptional and post-transcriptional activity [14]. The functions of circRNAs mentioned above are as gene expression regulation agents acting as transcriptional regulators. Some intronic circRNAs and ElciRNAs can regulate protein production by competing with the splicing of pre-mRNA and binding RNA polymerase II respectively [14]. However, the mechanisms of theses phenomena are still unclear [12]. As with linear RNAs, circRNAs can also bind to

RNA-binding proteins (RBPs) that are highly related to cancer, by affecting their functions as upregulated and downregulated activity. They may bind protein's active sites rather than other proteins that can act as competing elements. For instance, ciRS-7 can degrade the Argonaute 2 protein and inhibit transcription of mRNA [14]. Even though circRNAs can be thought as non-coding molecules, there is a significant chance of circRNA translation actually happening in cytoplasm [15]. Even if the actual translation mechanism remains uncertain, there are four types of mechanism for circRNA translation such as internal ribosomal entry site (IRES)-dependent mechanism, rolling circle amplification (RCA) mechanism, UTR of ribo-circRNA (cUTR)-dependent mechanism, M6A modification of RNA [16].

2.1 Types of CircRNAs

Despite the fact that formation mechanisms of circRNAs remain unclear, all types of CircRNAs are derived from pre-mRNA by using the canonical spliceosomal mechanism. They come into existence from exons of coding regions or from 3' UTR, 5' UTR, antisense RNAs, intergenic regions, and introns [17]. With the development of next generation sequencing technology, four types of circRNAs have been explored up-to-date. These are exonic circRNAs, which are composed from only one or more exons; intronic circRNAs, which consist of only introns; exonic-intronic circRNAs, which are derived from both exon and intron; and tRNA intronic circRNAs, that are derived from pre-tRNA intron [18]. However, human circRNAs mainly consist of exons, and almost 80% of identified circRNAs are also exonic circRNAs [17]. Although formation of circRNAs is not a clear issue, there are two known different models which are lariat splicing model and direct back-splicing model. In lariat back-splicing model, circulation of circRNAs happens after canonical splicing. [78] But in direct back-splicing models occurs before the canonical splicing. However, most exonic circRNAs can be generated by back-splicing formation at annotated exon boundaries or canonical splice signals by spliceosome. [42] Back-splicing is a type of alternative splicing event whose main

mechanism in which 5’ donor attacks 3’ upstream splice site and 3’-5’ phosphodiester bond has occurred; it is still an ambiguous discussion. [17] Also, the same primary transcript can generate linear mRNA, intronic circular RNA, exonic circular RNA as seen in Figure 1.

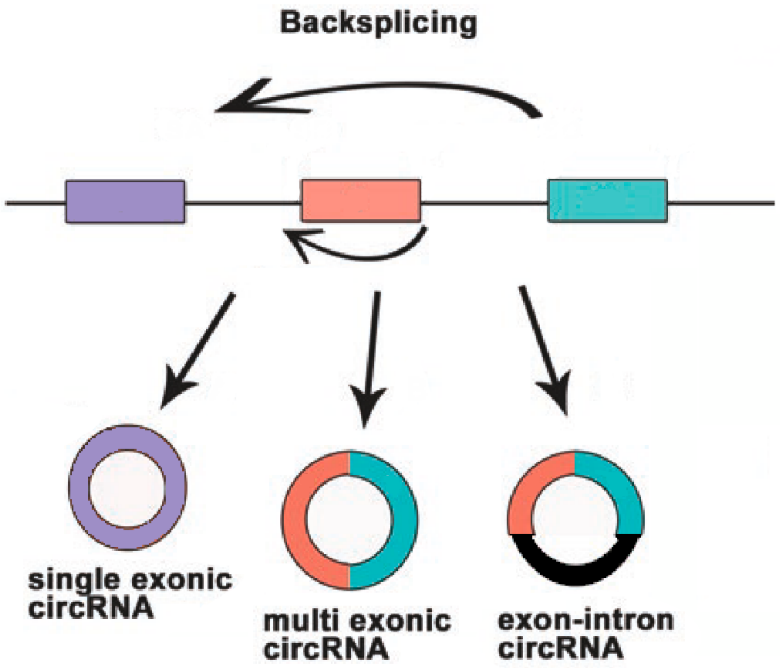

Figure 1 Formation of Exonic CircRNAs

2.2 CircRNAs in Cancer

As mentioned in the previous section, circRNA molecules are highly stable and abundant in the transcriptome of eukaryotic cells and have cell/tissue specificity and conservation properties [10]. Due to their circular structure, circRNAs are more resistant to exonuclease RNase R enzymes than linear RNAs, and therefore, they can be detected more readily in body fluids. Due to this property, circRNAs are a good candidate for clinical diagnostic and prognostic markers [19]. It is well known that microRNAs play an important role in many mRNAs activities and also it is directly related to solid tumor formation. The aforementioned miRNA sponges, as one of the circRNAs’ functions, are directly associated with cancer development [19].

Interaction between miRNA and circRNAs is not the only interaction of circRNAs that can be related to cancer development. Long non-coding RNAs, and RNA binding proteins may also bind to circRNAs. Alteration of glucose mechanism is one of the processes leading to cancer, and it can be regulated via some enzymes such as hexokinase (HK), 6-phosphfructa-1-kinase (PFK), and pyruvate kinase (PK). These enzymes work all together in the transformation of glucose into pyruvate and circRNAs can affect their activities. One of these circRNAs is circHIPK3 that is bountiful in pancreatic islets, decreased Slc2a2 expression that encodes GLUT2. Not only enzymes are affected via circRNAs, but some transcription factors also that are included in Warburg cycle also can be affected circRNAs such as HIF-1 by sponging miR-186, C-myc. Also, circRNAs play a part in glycolysis by regulation of signal pathways. Other metabolic processes affected by miRNA sponging are the lipid and amino acid metabolisms [20]. Moreover, accumulation of circRNAs may cause inhibition of its parental genes likewise downregulation of DNMT1 gene affected by accumulated circFECR1 [18].

2.2.1 Hepatocellular Carcinoma

At the early stage, hepatocellular carcinoma cannot be diagnosed because liver biopsies cannot be done routinely for screening due to the risk of complications [21]. HCCs contain high heterogeneity in terms of molecular phenotypes. This raises major challenges in clinical management. Tumor number, tumor size, vascular invasion, and performance status are used as metrics [4]. But these metrics are not enough necessarily enough to diagnose HCC properly and to select the treatment type, due to the phenotypic and genetic heterogeneity. The main reason for the high death rate in HCC is the lack of efficient treatment options. While Liver tranplantation or surgical resections can be efficient treatment options at the early stage of HCCs, as mentioned before, the percentage of the diagnosis of HCC at the early stage is only % 20 [22].

Self-sufficiency in growth signals, insensitivity to anti-growth signals, evading apoptosis, sustained angiogenesis, tissue invasion & metastasis and limitless repliction are the six hallmarks of cancer cells. Understanding the hallmarks of cancer leads researchers to understand the genes' connectivity on signaling pathways and discover new therapeutic agents [23]. Tyrosine kinase pathways, the Ras mitogen-activated protein kinase (Ras/Raf/MAPK), the phosphatidylinositol 3-kinase (PI3K)/AKT/mammalian target of rapamycin (mTOR), the Wnt/β-catenin signaling pathway, the ubiquitin/proteasome degradation and the hedgehog signaling pathway are the most crucial pathways in HCC [22]. Due to the high heterogeneity of HCC, the only approved drug is Sorafenib, which stops the signal of tumor blood formation and tumor growth by targeting multi-kinases in the advanced liver cancer [24]. Sorafenib targets the VEGFR, Ras/Raf/MEK/ERK, PDGFR, c-KIT, and RET [25].

2.2.2 CircRNAs Discovered in Relation to Hepatocellular Carcinoma

circC3P1, cSMARCA5, circMTO1 and circZKSCAN1 are circRNAs that are found in HCC tissues, which act as tumor suppressors [17]. CircC3P1 affects phosphoenolpyruvate carboxykinase 1 expression by sponging miR-461 in HCC cells [67], also cSMARCA5 downregulates miR-17-3p and miR-181-5p activities, increases the expression of TIMP metallopeptidase inhibitor 3 [68]. Also, circMTO1 sponges oncogenic miR-9 that is directly related to p21 expression [69]. circZKSCAN1 interferes with various cancer pathways and constrains cell proliferation, migration, and invasion [70]. Moreover, although its function remains unclear, hsa_circ_0004018 has related mechanisms in HCC. It associates with serum alpha-fetoprotein (AFP) level, tumor diameter [71]. circRNA_100338, circ_000839 sponges related miRNAs, miR-141-3p and miR-200b in HCC cells [72, 73]. In addition to these circRNAs, ciRS-7 is highly expressed in HCC cells directly linked to hepatic microvascular invasion and AFP level [11]. hsa_circ_0001649 stimulates tumor size and the formation of tumor embolus in HCC cells [75]. While

hsa_circ_0005075 promotes adhesion, Cdr1as stimulates cell proliferation which act as sponging miR-7 involved PI3K/Akt/ mTOR signaling pathway in HCC [76, 77]. Cdr1as may be thought as an HCC prognostic biomarker and a therapeutic agent for microvascular infiltration [11]. Also high expression of hsa_circ_100338 in HCC tissues indicates start of the metastasis progress [19]. A recent review of the literature on this topic [2017] found that after detecting 127 differentially expressed circRNAs and selecting the top five; circZFR, circFUT8, circIPO11 are the significantly distinguishable ones among 127 circRNAs and also they construct a circRNA-miRNA- mRNA network to emphasize the progression of hepatic cancer [26]. In their groundbreaking paper of 2017, Guo and Zhang studied circ-TCH whose expression level in HCC tissue is undoubtedly lower than normal tissue [27]. Moreover, circ_000839 upregulates RHoA by sponging miR-200b and the expression of circ_0067934 is high in HCC, also circ_0067934 sponges miR-1324 [28].

2.3 CircRNAs Databases

As mentioned in the earlier sections, circRNAs are involved gene regulation by sponging miRNAs and have an indirect relation to mRNA translation. Therefore, cirRNAs have become crucial hot points for transcriptomic studies. Several circRNA databases are available online including Circbase, TSCD, CircRNA (circRNADb), CIRCpedia V2, Deepbase V2, CircInteractome, CirclncRNAnet, CircBANK, circAtlas, CircR2Disease, CirCad, CircRNA Disease, Circ2Traits, CSCD, MiOncoCirc, and CircFunBase.

Among these databases, Circbase, TSCD and CircRNADb, CircR2Disease, CirCad, and CircRNA Disease are data collected from laboratory experiments. Circbase contains sequences of circRNAs and their genomic location via assembling 9 different articles related to detection of back-splice junctions for human, mouse, c. elegans, latimeria experimentally. They provide researchers with simple search, list search and table browser [29]. TSCD is known as Tissue-Specific-CircRNA-

Database which contains 302,843 tissue specific circRNAs in mice and humans by conducting RT-PCR experiments. Moreover, 6,795,157 miRNA and their circRNAs interactions in mouse, human and human fetal are included in TSCD. In addition to miRNA-circRNA interaction, it contains interactions of 37 RNA binding proteins information [30].

MiOncoCirc is the first clinical circRNAs database associated with cancer only. They conducted exome capture RNA-seq protocol for more than 2000 tumor samples. Besides discovering new circRNA not included in CircBase, they found 129,000 circRNAs and their isoforms [31].

CircRNADb consists of 31,914 human exonic circRNAs and contains related information for all the possible isoforms; 16,328 circRNAs having Open Reading Frames or IRES elements and 45 circRNAs from 37 genes related to their proteins expressed [32].

CircR2Disease is one of the manually curated databases, they collected all circRNAs in human, mouse, rat data up to March, 2018. Among 661 circRNAs, they assembled 725 associations with 100 diseases like Alzheimer disease, breast cancer, bladder cancer, cervical cancer, hepatocellular carcinoma, hepatoblastoma, liver cancer, and ovarian endometriosis [33]. Circad is also a manually curated database, containing 138 diseases related to 930 circRNAs in human, rat, mouse, chicken and suitable sequences of primers [34].

circRNA Disease is another manually curated database, they collected 800 published literature information including 330 circRNAs associated with 48 diseases up to 2017. Besides the disease names, they give researchers circRNAs' functional descriptions like "Overexpression of circ-104916 effectively inhibited the proliferation, migration and invasion abilities of GC cells." [35].

Circ2Traits uses 1,953 predicted circRNAs taken from an article [53]. Moreover, they identified 105 diseases associated with miRNA SNPs and argonaute interaction sites [36].

CircFunBase collected more than 7,000 functional circRNAs in a diverse range of species from articles published up to May 2018. By performing gene set enrichment analysis for cancer, they provide researchers with circRNA-miRNA interaction networks, circRNA expression patterns and circRNA-RBP interactions [37].

CIRCpedia, CircInteractome, CirclncRNAnet, circAtlas, Circbank are databases that were created computationally. Whereas CIRCpedia does not involve any interaction information, the others do.

CIRCpedia contains 267,782 circRNAs (18,943 circRNAs in human) and their gene expression levels including some disease samples by detecting back-splice junctions in human, mouse, rat, zebrafish, fly, worm from 180 RNA-seq datasets by using CIRCexplorer2 detection tool [38].

circAtlas uses the CIRI algorithm to detect circRNAs by taking normal tissue samples (184 RNA-seq libraries) in human, macaque and mouse. 283,384 circRNAs were discovered. Although it contains no interaction information, circAtlas is integrated into circad, CircR2Disease and circRNADisease databases [39]. Cancer-Specific-Circular RNAs (CSCD) is created by taking 82 cancer cell lines samples from ENCODE. CSCD detects 1,394,023 circRNAs using CIRI2, find_circ, circ_RNA, finder circexplorer algorithms. Also, there are millions of circRNA-miRNA relationships, 37 different RNA binding protein sites, and full-length sequences of open reading frame in this database [40].

CircInteractome is created for circRNAs interactions such as miRNA, RNA binding protein and identification of IRES elements by taking 65,526 circRNAs information from Circbase [41].

2.4 CircRNA Detection

CircRNAs have closed and covalent bonds between upstream a 3′ splice site and a downstream 5′ splice site of the pre-messenger RNA. Their general size is between 100 nt to larger than 4 kb, but in humans, a few hundred nucleotides. It is well known

that the formation of circRNAs is based on backsplice events occurring at exon boundaries [42]. With high throughput next generation sequencing technology, shotgun sequencing of many of short RNA reads have enabled researchers to detect circRNAs. Detection of circRNAs from RNA reads has some challenges in both experimental procedure and computationally. One of the reasons that circRNAs cannot be detected easily is that they have smaller sizes and mobility properties compared to linear isoforms. For instance, the amount of circRNAs is in between % 5 to 10 of their linear counterparts [12]. Library preparation is one of the important steps to detect circRNAs properly. Library preparation has following three important steps: RNA purification, size selection and RNA fragmentation, and the method of priming. The crucial point for detection of circRNA is that purification steps should be selected according to circRNAs structures. By the reason not having poly A tails, a poly(A) enrichment step would deplete circRNAs. The other recommended library preparation step is that a sample be treated with RNase R, due to their stability, they can be selected properly [42].

An additional challenge in circRNAs detection is computational challenges such as identification of back-splice junctions, mapper of choice and dependency on reference genome.

These challenges cannot be neglected due to the fact that these three challenges cause false-positives, uncertainty in linear/circular isoform ratio, and low sensitivity or accuracy [42].

Presence of gene annotation in circRNA detection algorithms makes it more sensitive and accurate in comparison to pseudo-reference algorithms or de novo algorithms since gene annotation enables the algorithm to select exon boundaries accurately. Pseudo-reference algorithms or de novo algorithms use non-canonical signals at the cost of high false positives [43]. Most of these algorithms choose GU-AG splicing sequences which are very small as non-canonical signals. However, there is one disadvantage in choosing gene annotation dependent algorithms. They may create blind spots since they do not use non-canonical splice signals [12].

Another approach is the split-alignment method in which identifies back-splice junctions from mapping information [42]. As mentioned earlier, there are two main approaches to detect circRNAs. One of them is a split-alignment approach in which back-splice junctions are split into segments and then aligned to a reference genome. UROBORUS, DCC, CIRCexplorer2, circRNA_finder, MapSplice, Segemehl, find_circ and CIRI are the circRNAs detection tools based on split-alignment approach. Whereas UROBUS, DCC and CIRCexplorer2 prefer splice-aware mappers like STAR, Tophat; circRNA-finder, MapSplice, Segemehl, find_circ and CIRI use non-splice-aware mappers like Bowtie, BWA [43]. Figure 2 represents the categorized according to different methods.

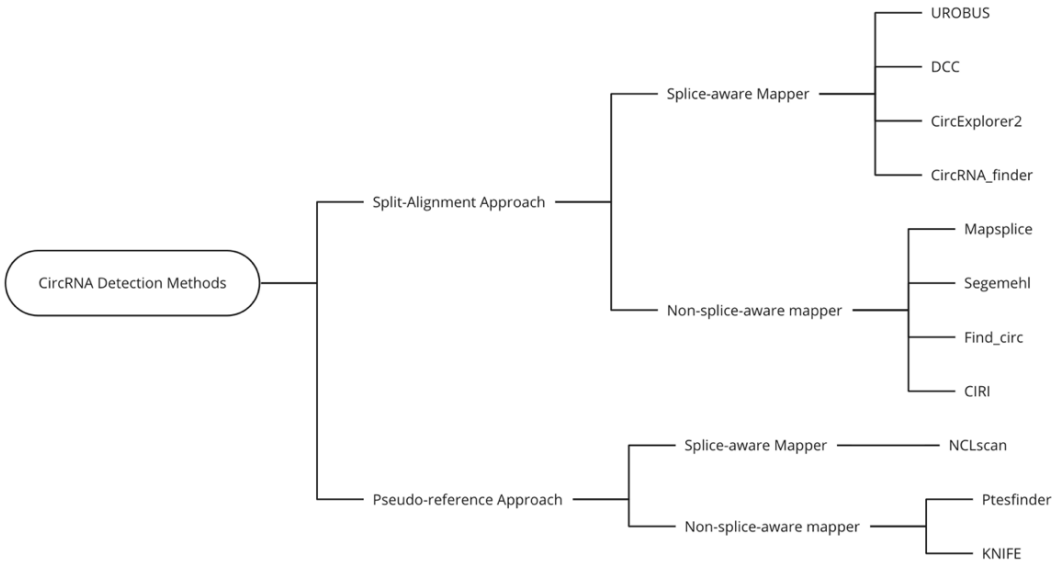

Figure 2 CircRNAs Detection Methods

CIRCexplorer2 provides researchers with a two-phase approach. At the first phase, RNA-seq data should be mapped via TopHat. Tophat-Fusion should be applied for unmapped reads. In addition to TopHat/Tophat-Fusion, STAR aligner is also applicable for this workflow. De novo assembly for single-end data and pair-end data is also suitable for researchers [44]. A new approach is described in DCC [62] and circRNA_finder [63] due to usage of STAR alignment. STAR aligner can directly

access potential chimeric reads (emblematic for circRNAs) instead of chaining multiple subsequent mapping steps [45]. CIRI can work with only SAM files with the BWA aligner. In contrast to other algorithms, they use a novel algorithm taking into account potential paired chiastic clipping signals (PCC) detection in the sequence alignment/SAM to remove false positives [46]. The MapSplice detection tool also uses a two phase approach similar to CIRCexplorer2. In the tag alignment step, mRNA tags are mapped to a reference genome, then the identification of candidate tag alignments is performed. The second step is the splice inference phase in which splice junctions that appear in the alignments of each tag to infer a splice significance value [47]. KNIFE is one of the pseudo-reference algorithms taking into account all possible exon-exon boundaries from gene annotation file before the alignment step. It calculates a statistical score for each read, based on alignment properties, and number of mismatches [48]. Among all detection tools, KNIFE, CIRI and CIRCexplorer2 have higher accuracy and sensitivity [12].

2.5 Downstream Analysis of CircRNAs

For accurate circRNA quantification and differential expression analysis, there should be a direct comparison between circular and linear RNA expression. This is because the sequences of circRNAs produced from backsplice events of exons may overlap with the sequences from linear isoforms transcribed from the same gene loci. FUCHS, CIRI-AS, CircView, Sailfish-cir, CirPro and CircTest are the computational approaches for downstream analysis of circRNAs. Sailfish-cir and CircTest can be used for downstream analysis for expression analysis and quantification of circRNAs. FUCHS, CIRI-AS, and CircView can be used to perform miRNA seed analysis, detect internal structure, and estimate the potential protein-coding, respectively [49] CIRI-full algorithm provides researchers with accurate quantification and differential analysis at the isoform level [43]. CLEAR is another tool that provides researchers with direct comparison giving a linear to circular expression level ratio. CLEAR consists of mainly two steps; alignment and

quantification. At the alignment step, RNA-seq fragments are mapped by HISAT2 against GRCh37/hg19 human reference genome with known gene annotations for linear RNA quantification analysis. Then, HISAT2-unmapped fragments are then mapped to the same GRCh37/hg19 reference genome using TopHat-Fusion for circRNA quantification [50].

CHAPTER 3

METHODOLOGY

3.1 Overview

CirComPara is one of the integrative circRNA detection and quantification tools. It starts with the quality control analysis by using FastQC and Trimmomatic. The analysis continues with alignment to reference genome using HISAT2. Later, the reads that could not be mapped to reference genome are used for the detection of circRNAs. The other mapped reads are used for the linear transcriptome reconstruction and then linear gene expression, by using Cufflinks and Cuffnorm respectively. After getting unmapped reads, they are ready to be inputs for detection methods. Finally, by evaluating the ratio between circular and linear isoform expression, a specific score named the circ_score is calculated

3.2 Data Sources

In this work, RNA sequences of two different cell lines treated with Sorafenib and PI3K inhibitors were used as fastq file format. These data were taken from CANSYL laboratory in METU. The raw data is available in the NCBI SRA database under BioProject accession number PRJNA556552 [54].

HUH-7 and Mahlavu cell lines were treated with Sorafenib, PI3K-α, PI3K-β, Sorafenib with PI3K-α, Sorafenib with PI3K-β as their combinations in IC50 concentration for 48 hours. HUH-7 cell line is a well-differentiated hepatocellular carcinoma cell line which was taken from male and [55] It is generally used to investigate HCC and HCV. [57] Mahlavu cell line is a hepatocellular carcinoma cell line, but it is poorly differentiated and has low cytoplasm. [56, 58] While expression

of HUH-7 cell line can be high in the early stage of HCC, Mahlavu cell line can be found in the advanced levels of HCC. [59]

Sorafenib is one of the multiple targeted drugs for the advanced level of HCC. By targeting the platelet-derived growth factor receptor, vascular endothelial growth factor receptor and hepatocyte factor receptor in Raf-1, B-Raf and Ras/Raf/MEK/ERK signaling pathways, it suppresses the tumor cell proliferation. [60]PI3K/Akt/mTOR pathway is one of the crucial pathways associated with many cancer including HCC. [64] PI3K kinases are from family of lipid kinases and categorize into three classes: class I, II, and III In class I, they have four isoforms such as PI3K-α, PI3K-β and PI3Kδ, and PI3Kγ. These isoforms take a role in PI3K pathways by dephosphorylating of PIP3 to PIP2 PTEN is a tumor suppressor protein getting involved in this pathway by regulating PI3Ks' activities. [65]. While PI3K-α isoform regulates the cellular growth, metabolism and angiogenesis, PI3K-β are related to inflammatort cells by G protein-coupled receptors. [54] In Table 1, all the information and abbreviations of drug name and conditions are given.

Table 1 All the drug conditions used in this research

Drug Name	Conditions
DMSO	DMSO
Sorafenib	SOR
PI3K-α inhibitor	ALPHA
PI3K-β inhibitor	BETA
Sorafenib with PI3K-α inhibitor	S_ALPHA
Sorafenib with PI3K-β inhibitor	S_BETA

3.3 Data Pre-processing

Data pre-processing of RNA sequences contains three main steps, which are the quality control step, trimming, and extraction of unmapped reads. The quality control step was conducted by using FastQC program [5]. It enables researchers to have a html file including per base sequence quality, per tile sequence quality, per sequence quality scores, per sequence GC content and adapter content. After taking the results of FastQC program, Trimmomatic [6] is used to trim RNA sequences, according to given parameters such as sliding window quality filtering, maximum information quality filtering, leading, trailing, minimum length and average quality. HISAT2 [7] is one of the fast and accurate alignment programs for not only transcriptome but also whole-genome data. It is used to get mapped and unmapped reads from RNA sequences for this work. Later while mapped reads will be used to get linear expression data, unmapped reads will be used to detect back-spliced junctions by using different detection methods.

3.4 Detection of Circular RNAs in HUH7 and Mahlavu Cell Lines from RNA-seq data

After getting unmapped reads from HISAT2 alignment program, the unmapped reads are used as input data for all the detection methods including findcirc, circRNA_finder, DCC, CIRI, Circexplorer2 with Tophat2 and Circexplorer2 with Star. All the detection methods need the index files including HISAT2 indexes, bwa indexes, bowtie2 and bowtie indexes, STAR indexes. While findcirc, CIRI and circRNA_finder are the annotation independent detection methods, DCC and Circexplorer2 are the annotation dependent detections methods. All detection methods including findcirc, circRNA_finder, DCC, CIRI, Circexplorer2, need to re-map the unmapped sequences by using different aligners to detect the back-spliced junctions.

Circexplorer2 is in splice-aware category under the annotation-dependent methods. It can use either Tophat2 or STAR as aligner. It requires unmapped sequencing reads. It also needs gene and exon annotations to detect the back spliced junctions points. Therefore genome annotation and the whole genome are the mandatory files. If one uses Tophat2 as an aligner, bowtie and bowtie2 indexes are also needed. If STAR is used as an aligner, genome indexes which are generated by using the STAR program and also annotation and whole genome files are also need. CIRI uses bwa-MEM as aligner to re-map unmapped sequences. Also, it needs gene annotation files and GT-AG signals. DCC uses STAR as aligner, like Circexplorer2. So, it also needs genome annotation and whole genome. To detect circRNAs, it also uses GT-AG signals. circRNA_finder uses annotated exons within gene and GT-AG signals by using Tophat2 as aligner. And it also uses whole genome, bowtie indexes and gene annotation files. Findcirc uses Bowtie to remap unmapped sequences which is coming from HISAT2. It uses also bowtie2 index and whole genome files to detect the back spliced junctions knowing GT-AG signals.

After installation of the prerequisite software packages (TopHat2, Tophat_Fusion, STAR, BWA, pysam, pybedtool, Bowtie, Bowtie2, Samtools, Bedtools, ggplot2, Python, R, Scons) and all the detection methods (findcirc, circRNA_finder, DCC, CIRI, Circexplorer2), CirComPara uses two main files including comma separated file; meta.csv and a Python file; vars.py for execution. Meta file is a comma separated files which contains pathways of sample's and adaptors' fastq files an, sample names, conditions. The title of sample gives what type of condition were applied for each cell line and condition represents the DMSO (control) and drug treatment ways. The other main file is the vars python file, it contains all the parameters including types of detection methods with their parameters, parameters for aligners, and pathway of all annotation and index files as seen Table 2. CirCompara allows the researchers to generate indexes like HISAT2 indexes, BWA indexes, Bowtie2 and Bowtie indexes and STAR indexes. In this work, they have been generated just before to execute the CirComPara. GRCh37 whole genome fasta file and gene annotation file have been downloaded from Ensembl. In this work, the

different pipelines have been worked for HUH-7 cell line and Mahlavu cell line separately.

Table 2 CirComPara Execution File

Meta	Name of Meta File
Genome Fasta	Absolute Path of GRCh37 whole genome fasta
Annotation	Absolute Path of GRCh37 Genome Annotation file (gtf extension)
CPUs	Available CPU in the computer
Preprocessor	Preprocessor Name (Trimmomatic)
Preprocessor Parameters	Preprocessor Parameters
CircRNA Methods	Types of Detection Methods
Toogle Transcriptome Reconstruction	Information related to whether Annotation Files used or not
Min Reads	Number of reads to take account as a circRNAs expressed
HISAT2 Extra Parameters	Single-end or Paired End Reads
BWA Parameters	BWA Parameters
Tophat Parameters	Tophat2 Parameters
STAR Parameters	STAR Parameters
Genome Index	Absolute Path of HISAT2 Index Files
BWA Index	Absolute Path of BWA Index Files
Bowtie2 Index	Absolute Path of Bowtie2 Index Files

Table 2 (Cont'd)

STAR Index	Absolute Path of STAR Index Files
Gene Pred File	Absolute Path of genePhred File
Linear Counter	Selection of Method for Linear Expression
DCC Extra Parameters	DCC Parameters
Find_circ Extra Parameters	Find_circ Parameters
Sam Sort MM	Maximum required memory per thread
BYPASS	Whether you want to circular to linear expression ratio or not

In Table 2, trimmomatic follows six parameters: (1) maximum information filter to balance the benefits of retaining longer reads against the costs of retaining bases with error; the read length and strictness are selected 40 and 0.5 respectively. (2,3) Leading (20) and Trailing (20) to remove low quality bases from beginning and end respectively. (4) Sliding Window to trim the bases considering and the number of bases average across and the average quality as 4:30. (5) Minimum Length to cut the reads whose length is below the stated length as 35. (6) Average Quality to drop the reads if their quality is below the specified leves as 30. [6] The toogle transcriptome reconstruction should be False because the annotation files are used for DCC, CircExplorer2 and circRNA_finder. HISAT2 extra parameter determined the reads are whether single end or paired end. All RNA sequences that are used for this work are single-end.

CIRI uses BWA as one of the filtering options to have read length in the junctions and quantity/quality of reads. [46] CirComPara allows the users to change the mapping quality score and the number of removing maximum exact matches

according to specified number of BWA by changing BWA parameters ('-T', '19', '-c', '1') as seen in Table 2. To get uniquely mapped reads, one of the Tophat parameter *(--max-multihits*) was set as 1.

Execution file vars also include Fix Read Header parameter, it is selected as True to trim Fastq header to the first white spaces. By this way, Fastq headers contain only reads ids.

STAR [61] is one of the RNA-seq aligners which can also detect non-linear transcripts so DCC and CircExplorer2 are used to get chimeric read alignments. To filter output file, STAR is used --outSJfilter* and –outFilter* parameters such as *outFilterMultimapNmax* (the maximum number of multiple alignments allowed for a read, 1), *outSJfilterOverhangMin* (minimum length for splice juntions for motifs; non-canonical motifs, GT/AG and CT/AC motif, GC/AG and CT/GC motif; and AT/AC and GT/AT motif, 15, 15, 15, 15) , *outFilterScoreMin* (if alignment score is lower than the specified score, it will discard, 1), *outFilterMatchNmin* (if the number of matched bases is lower than the specified score, it will discard, 1)and *outFilterMismatchNmax* (alignment will be output only if it has no more mismatches than specified value, 2). [61] For alignments and seeding options, STAR allows the users to change so many features like *alignSJoverhangMin (*minimum overhang for unannotated junctions,15), *alignSJDBoverhangMin* (minimum overhang for annotated junctions, 15), *seedSearchStartLmax* (the search start point through the read, 30). [61] *chimSegmentMin* (the minimum mapped length of two segments, 15)*, chimScoreMin* (minimum score of the chimeric segments, 15)*, chimScoreSeparation* (the difference between the best chimeric score and the next one, 10)*, chimJunctionOverhangMin* (minimum overhang for a chimeric junction, 15) are the parameters that are used to get chimeric read alignments. [61] CirComPara provide researchers with estimation circRNA-host gene and linear expression with lin_counter parameter as in seen in Table 2**.** DCC can also make an estimation circRNA-host gene, so the user can choose either dcc or ccp. In this work, ccp was chosen to get also linear expression data. In Table 2, five extra parameters were

added in DCC pipeline: (1) *fg*; not allow to span circRNAs canditates more than one gene (2) *M*; to distinguish circRNAs canditates from mitochondrial chromosomes. (3) *Nr*; minimum count in one replicate, 1 and number of replicates the candidate has to be detected in, 1. (4) *F*; to filter the circRNA candidate regions (5) *ss*; specification for single end sequences

Find_circ enables the user to select alignment score margin of the best anchor alignment(*--best-qual*), it is selected as 40 for this study.

All the details of meta and vars file can be founded in Appendix A.

CHAPTER 4

RESULTS

4.1 Quality Control Step and Read Processing

CircRNA analysis was conducted with two different cell lines; HUH-7 and Mahlavu treated with PI3K-α, PI3K-β, Sorafenib, DMSO, Sorafenib and PI3K-α, Sorafenib and PI3K-β.

The quality control analyses of RNA sequences of HUH7 and Mahlavu cell lines are done with FastQC. As an example, the per base sequence quality graph and per sequence quality score graph of RNA sequences of HUH7 cell line treated with PI3K-α is given in the Figure 3 and 4 respectively. All the remaining quality results can be found in Appendix A.

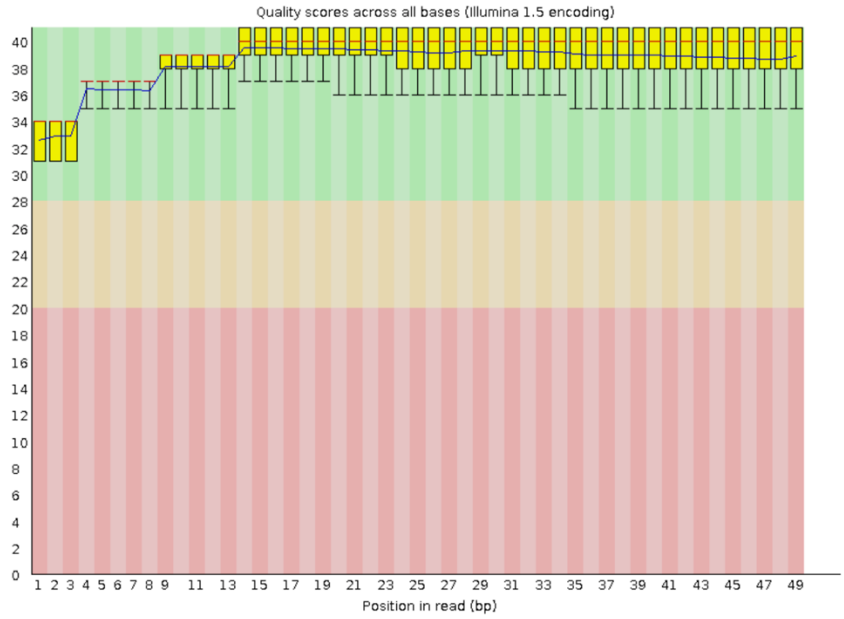

Figure 3 Per Base Sequence Quality of HUH7 cell line treated with PI3K-α

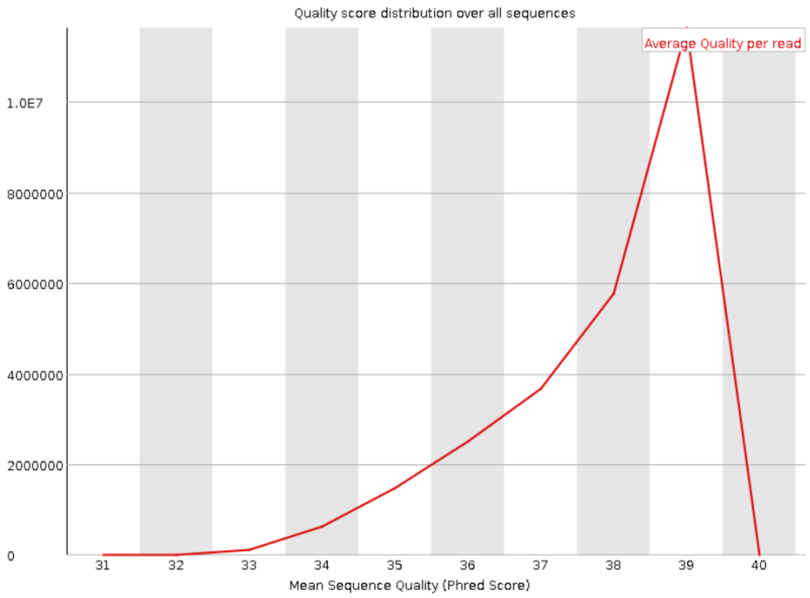

Figure 4 Per Sequence Quality Scores of HUH7 cell line treated with PI3K-α

After taking all the quality results, by using the trimmomatic program, some of the reads were dropped according to the specified parameters. After trimmomatic processing, clean reads are mapped through the linear genome by using HISAT2. The results of the mapping steps are given in the Table 3.

Table 3 Per Sample Read Count in FastQC, Trimmomatic and HISAT2

Sample	Raw reads	Clean reads	Dropped	Linearly Mapped	Linearly Unmapped
H_cell_ALPHA	33192992	25864157	7328835	23624132	2240025
H_cell_BETA	32730853	25286279	7444574	23754652	1531627
H_cell_DMSO	37827459	29572209	8255250	27836626	1735583

Table 3 (Cont'd)

H_cell_S_ALPHA	34017394	26486274	7531120	24626895	1859379
H_cell_S_BETA	28579376	22373643	6205733	21081585	1292058
H_cell_SOR	36506606	28448087	8058519	26846788	1601299
M_cell_ALPHA	34302640	26369733	7932907	24630825	1738908
M_cell_BETA	29776671	22884625	6892046	21503468	1381157
M_cell_DMSO	33421422	25801151	7620271	24228792	1572359
M_cell_SOR	32116151	24659599	7456552	23195895	1463704
M_cell_S_ALPHA	31897951	24272549	7625402	22518000	1754549
M_cell_S_BETA	37221329	28448492	8772837	26798516	1649976

Figure 5 and 6 include all the comparisons of the proportions of the reads coming from the processing steps of HUH7 cell lines and Mahlavu cell lines with different treatment options in each sample respectively.

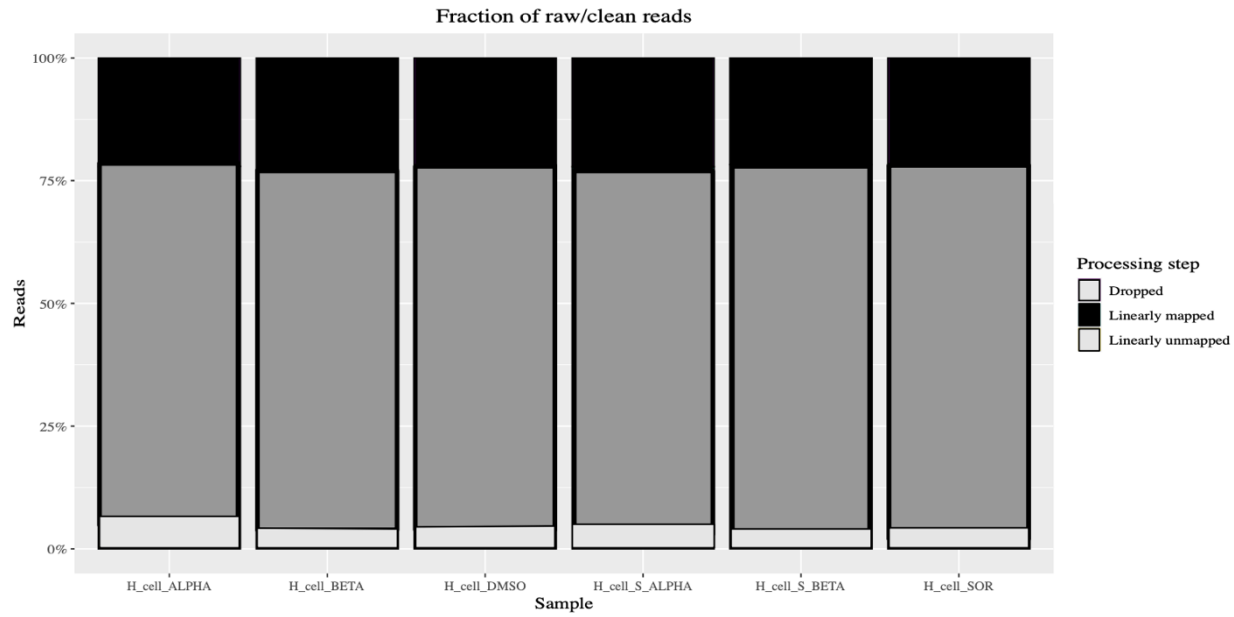

Figure 5 Proportions of the reads resulted from the processing steps of HUH7 cell line

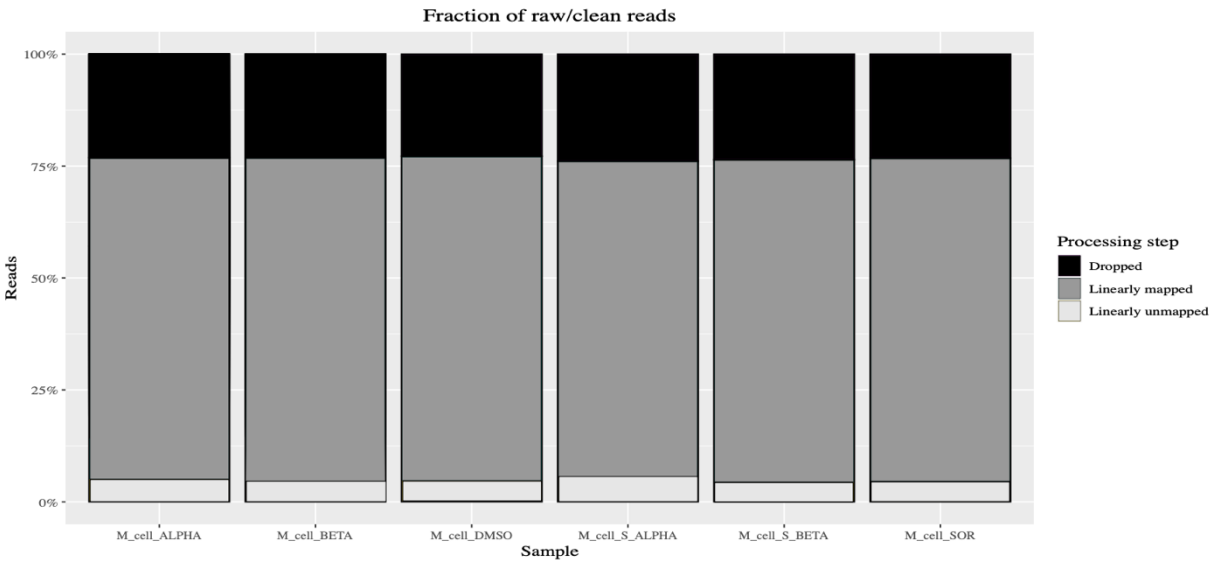

Figure 6 Proportions of the reads resulting from the processing steps of Mahlavu cell line

4.2 CircRNAs Detection

After read processing steps of all the RNA sequences including HUH-7 and Mahlavu cell line with different treatment options, circRNAs detection steps have been carried out by using six different methods; findcirc, circRNA_finder, DCC, CIRI, circexplorer2 with tophat and circexplorer2 with star. In total 11,245 and 127,83 circRNAs have been detected in HUH-7 cell lines and Mahlavu cell lines, respectively. To get only reliable circRNAs, two filter options have been applied, one of the filters was to have more than two reads in the sequences, and the other filter is that circRNAs must be detected with least two detection methods. After applying the first filter step, there are only 857 and 850 circRNAs having more than 2 reads in the HUH-7 and Mahlavu cell lines respectively. Then, 136 in HUH-7 and 122 circRNAs in Mahlavu have been labelled as reliable circRNAs, by taking into account the second filtering step. Figure 7 shows all the numbers of circRNAs detected in HUH-7 cell lines with respect to detection methods and all reliability information.

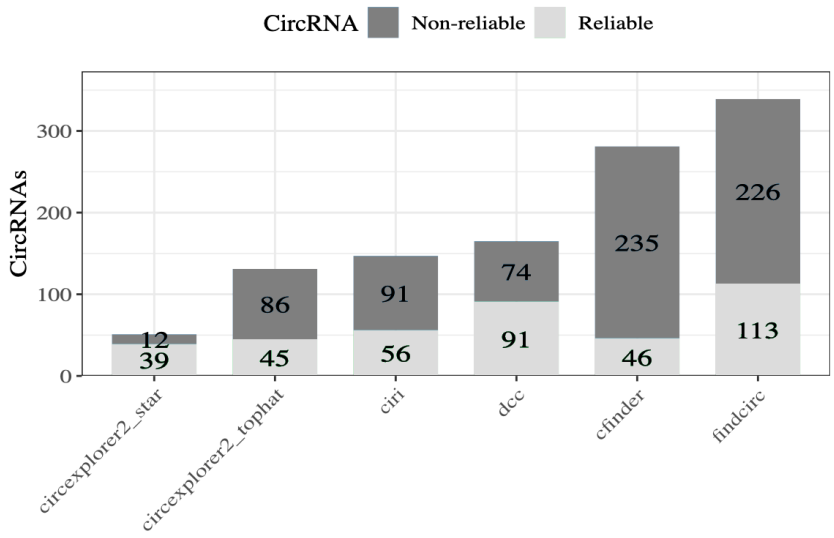

Figure 7 Total Detected CircRNAs in HUH-7 cell lines with different detected methods

All non-realiable and reliable circRNAs detected in Mahlavu cell lines according to different detection methods are shown in Figure 8.

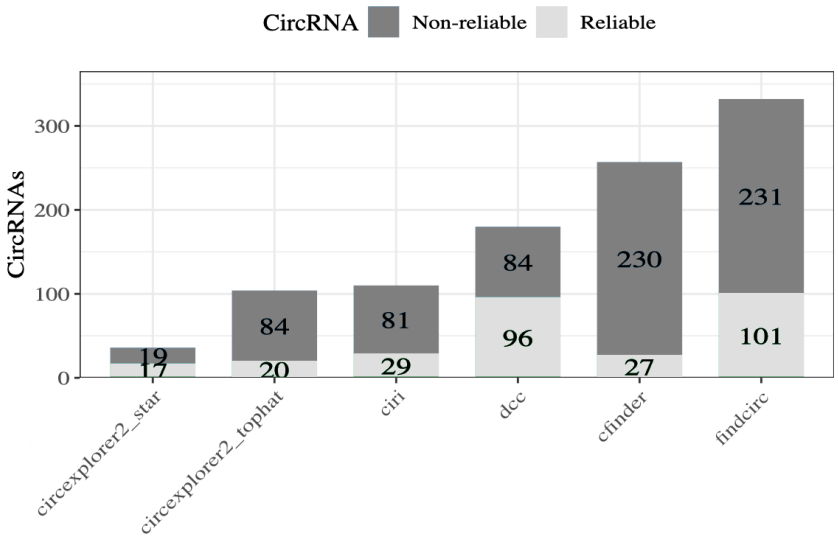

Figure 8 Total Detected CircRNAs in Mahlavu cell lines with different detected methods

Also, Table 4 contains all the related information with respect to applied treatment options with HUH-7 cell lines and Mahlavu cell lines.

Table 4 All the detected circRNAs with respect to different cell lines and treatment options

Sample	>= 2 reads	>= 2 reads & >= 2 methods
H_cell_S_BETA	259	62
H_cell_SOR	315	64
H_cell_ALPHA	330	65
H_cell_S_ALPHA	308	66
H_cell_BETA	296	56
H_cell_DMSO	359	69

Table 4 (Cont'd)

M_cell_ALPHA	340	51
M_cell_BETA	282	47
M_cell_DMSO	311	55
M_cell_SOR	308	52
M_cell_S_ALPHA	307	45
M_cell_S_BETA	343	65

Among all the types of circRNAs, in HUH-7 cell lines, there are 97 reliable exonic circRNAs overlapping 97 genes, in Mahlavu cell lines 79 reliable circRNAs overlap 84 genes.

The further analysis has been done by analyzing only HUH-7 cell lines with different treatment options and only Mahlavu cell line with different treatment options and then by comparing HUH-7 and Mahlavu cell lines.

These numbers contain all types of circRNAs including exonic, intronic and intergenic. But the analyses will continue only with the exonic circRNAs. The examination has been done through into four different approaches by looking at the circRNAs disappearance changes into circular expression through circular to linear expression ratio (CLR). These are the CLR score which increased after the treatment, CLR score which decreased after the treatment, some circRNAs have disappeared after the treatment and some circRNAs have emerged after the treatment.

After the detection of circRNAs, the investigation was continued by finding downregulated/ upregulated circRNAs and parental genes. In comparison between gene expression of DMSO and all the therapeutic agents including PI3K-α inhibitor, PI3K-β inhibitor, Sorafenib, Sorafenib and PI3K-α inhibitor, Sorafenib and PI3K-α inhibition, downregulated/upregulated genes were determined by considering FPKM

values in Cytoscape. We tried to filter circRNAs whose CLR score increased after the treatment, and which have emerged after the treatment.

4.2.1 HUH-7 Cell Line with Different Treatment Options

4.2.1.1 HUH-7 Cell Line treated with PI3K-α inhibition

While 65 reliable CircRNAs have been detected in HUH7 cell line treated with PI3K-α, 69 reliable circRNAs have been detected in HUH7 treated with only DMSO. All among circRNAs, there are 44 and 49 exonic circRNAs detected in HUH-7 cell line treated with PI3K-α and DMSO respectively.

There are 15 exonic circRNAs that are only treated with PI3K-alpha. Among all 15 circRNAs, 9 of them were annotated by circAtlas, circBase, circRNAdb and circpedia2 as seen in Table 5.

Table 5 CircRNAs in HUH-7 cell line treated with PI3K-α inhibition that doesn't belong to the before the treatment group. (only PI3K-α, not containing DMSO)

CircRNA ID	Gene Name / Annotation
10:70719562-70720005	DDX21 (hsa-DDX21_0005)
12:114392936-114393010	RBM19 (hsa_circ_30158)
15:44843074-44843720	EIF3J (hsa-EIF3J_0001)
19:18047219-18047388	CCDC124 (hsa-CCDC124_0001)

Table 5 (Cont'd)

19:39216523-39217653	ACTN4
1:117944808-117957453	MAN1A2 (hsa-MAN1A2_0002)
22:31476436-31588688	RNF185|SMTN
2:171819381-171822314	GORASP2
3:196210641-196215554	RNF168 (hsa-RNF168_0002)
4:4276094-4276496	LYAR
6:18256592-18258636	DEK (hsa-DEK_0018)
8:62563609-62563683	ASPH
8:68044186-68049838	CSPP1 (hsa-CSPP1_0011)
MT:12360-12563	MT-ND5

20 exonic circRNAs have disappeared after the treatment (only DMSO, not containing PI3K-α) as seen in Table 6. 11 of them are annotated circRNAs in circRNA databases.

Table 6 CircRNAs in HUH-7 cell line with DMSO (not containing PI3K-α)

CircRNA ID	Gene Name / Annotation
10:70496711-70496805	CCAR1
10:95353711-95360223	RBP4

Table 6 (Cont'd)

10:95353731-95360223	RBP4
12:103232126-103232341	PAH
15:45250571-45315547	C15orf43|SORD
16:418328-418560	MRPL28 (hsa-MRPL28_0001)
1:155571864-155695810	DAP3|DAP3P1 (hsa-RP11-243J18_0001)
1:212148508-212161345	INTS7 (hsa-INTS7_0014)
1:228286421-228286476	ARF1
1:35652831-35653691	SFPQ (hsa-SFPQ_0037)
1:92446200-92446938	BRDT (hsa-BRDT_0006)
20:30732940-30733153	TM9SF4
22:42912017-42973104	RRP7A|RRP7B (hsa-RNU6-513P_0001)
2:113057426-113057606	ZC3H6 (hsa-ZC3H6_0005)
2:136505837-136530117	UBXN4 (hsa-UBXN4_0022)

Table 6 (Cont'd)

3:169694734-169706147	SEC62 (hsa-SEC62_0004)
7:102962379-102963241	DNAJC2 (hsa-DNAJC2_0003)
7:158662546-158669382	WDR60 (hsa-WDR60_0005)
8:86050270-86050710	LRRCC1
MT:13653-13854	MT-ND5

29 exonic circRNAs can be detected in both HUH-7 cell line treated with PI3K-α and DMSO as seen in Table 7. This table also contains circular isoform expression changes in these circRNAs according to circ_score.

Table 7 Common detected circRNAs in HUH-7 cell line with both PI3K-α and DMSO

CircRNA ID	Gene Name / Annotation	Circular Expression Changes
12:103306569-103306676	PAH	decrease
12:103306569-103311003	PAH	decrease
12:57108191-57108418	NACA	increase
15:53815415-53815519	WDR72	increase

Table 7 (Cont'd)

15:64448066-64448116	SNX22 (hsa-VMP1_0017)	increase
17:57886157-57889147	VMP1	decrease
17:79539043-79589304	NPLOC4	increase
17:79801680-79801750	RP11-498C9.2	increase
18:56008910-56024484	NEDD4L (hsa-MAN1A2_0004)	decrease
1:117944808-117948267	MAN1A2 (hsa-MAN1A2_0003)	increase
1:117944808-117963271	MAN1A2 (hsa-RP1-283E3_0030)	increase
20:32879294-32880197	AHCY	decrease
20:57485407-57485456	GNAS (hsa-SMARCC1_0045)	decrease
3:47663697-47680270	SMARCC1	decrease
3:49397297-49397427	RHOA	increase
4:74285971-74286038	ALB	increase
5:40832650-40834308	RPL37 (hsa-DEK_0003)	increase
6:18236683-18258636	DEK	increase

Table 7 (Cont'd)

7:100482834-100482991	SRRT	increase
7:100482837-100482991	SRRT	increase
7:97946549-97946668	BAIAP2L1	increase
8:17503423-17503518	MTUS1	increase
8:27462605-27462666	CLU	decrease
8:61763591-61763663	CHD7	increase
MT:10924-11206	MT-ND4	decrease
MT:1690-1894	MT-RNR2	decrease
MT:8927-9189	MT-ATP6	decrease
X:133087077-133087238	GPC3	decrease

RPS12, CCDC124, MT-ATP6, MT-NT4, MT-NT5, NEDD4L, ACTN4, ALB, CLU, GPC3, SRRT are upregulated genes according to linear expression data in RNA sequences of HUH-7 cell line treated with PI3K-α inhibition. Among 11 genes, there are only 4 genes (ALB, RPS12, NEDD4L and CLU) which have possible and known miRNA interactions. While linear expression information was calculated through CirComPara, miRNA-gene interactions downloaded from miRtargetlink 2.0. However, only ALB could be selected as both having miRNA interaction and increased circRNAs expression. CircRNA has found in 4:74285971-74286038 location on ALB and it is not annotated in circRNA databases so there is not any known miRNAs interaction. Moreover, this circRNA has 0.0224 CLR score, thus it does not consider as highly expressed circular isoform with respect to linear isoform. All linear isoforms of circRNAs were taken and put the Enrichr to get gene ontology analysis. But there were no significant biological process results, other than the

positive regulation of endocytosis according to adjusted p value. GPC3, NEDD4L, CLU are the genes involving the positive regulation of endocytosis.

There are differentially expressed genes circRNAs which have upregulated and downregulated linear isoforms and potential miRNA interaction. These linear isoforms have their own relations with related circRNAs as seen in Figure 9.

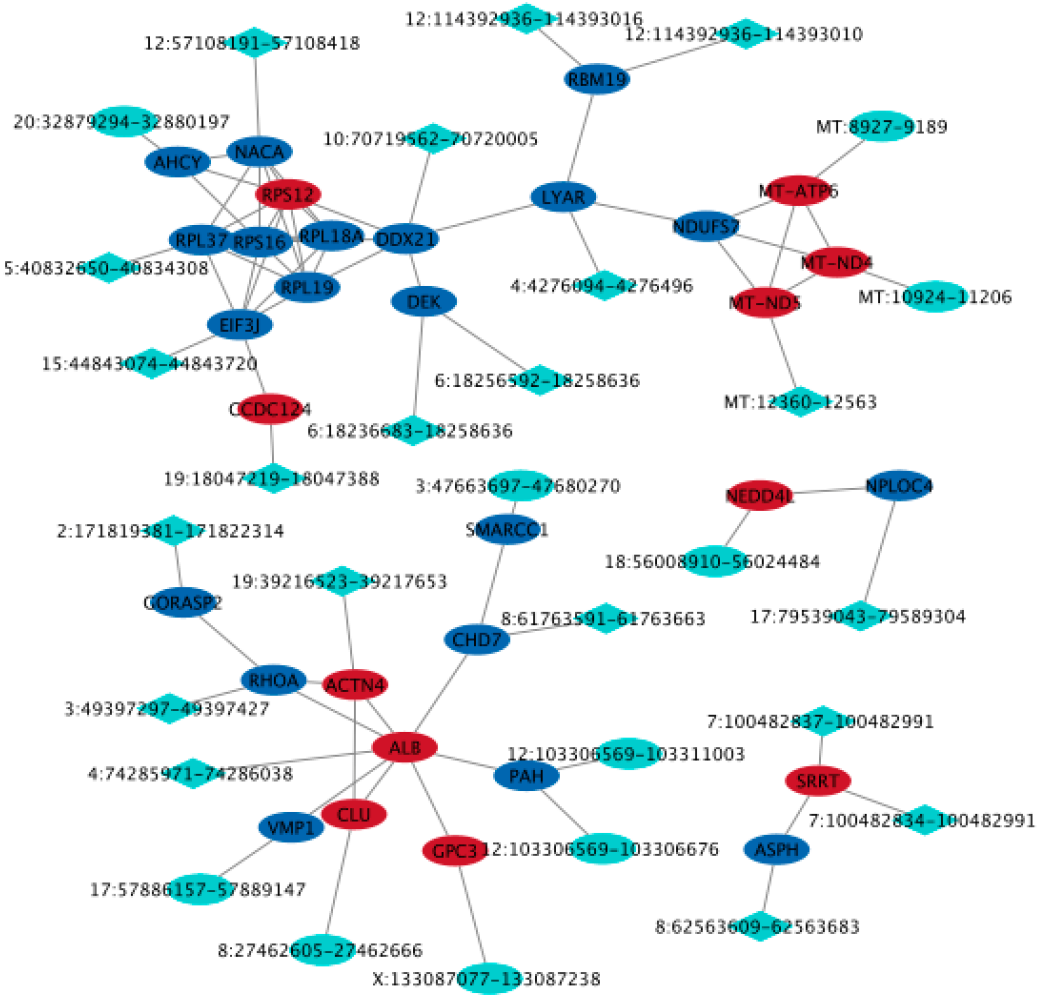

Figure 9 Differentially Expressed CircRNAs and Linear Isoforms Relations in HUH-7 Cell Line Treated with PI3K-α Inhibition

4.2.1.2 HUH-7 Cell Line treated with PI3K- β inhibition

While 56 reliable circRNAs were detected in HUH-7 cell line treated with PI3K-β inhibition, 40 of them are exonic circRNAs. As seen in Table 8, there are 10 exonic circRNAs detected in HUH-7 cell line treated with PI3K-β inhibition. (That doesn't contain DMSO) 5 of them are annotated by circRNAs databases.

Table 8 CircRNAs in HUH-7 cell line treated with PI3K- β inhibition that doesn't belong to the before the treatment group (only PI3K-β, not containing DMSO)

CircRNA ID	Gene Name / Annotation
13:46577274-46619651	ZC3H13 (hsa-ZC3H13_0004)
15:44843074-44843720	EIF3J (hsa-EIF3J_0001)
17:76000251-76027923	TNRC6C (hsa_circ_0008622)
1:117944808-117957453	MAN1A2 (hsa_circ_0000117)
22:31476436-31588688	RNF185|SMTN
3:145820542-145924548	PLOD2|PLSCR4
4:88005272-88012981	AFF1 (hsa_circ_0070382)
7:90034833-90034922	CLDN12|CTB-13L3.1

Table 8 (Cont'd)

9:2837180-2837276	KIAA0020
MT:12360-12563	MT-ND5

There are 19 exonic circRNAs that were detected in HUH-7 cell lines which were treated only DMSO and 8 of them are annotated in circRNAs databases as seen in Table 9.

Table 9 CircRNAs in HUH-7 cell line with DMSO (not containing PI3K-β)

CircRNA ID	Gene Name / Annotation
10:70496711-70496805	CCAR1
10:95353731-95360223	RBP4
12:103232126-103232341	PAH
15:45250571-45315547	C15orf43|SORD
16:418328-418560	MRPL28 (hsa-MRPL28_0001)
1:117944808-117948267	MAN1A2 (hsa-MAN1A2_0004)
1:155571864-155695810	DAP3|DAP3P1 (hsa-RP11-243J18_0001)
1:228286421-228286476	ARF1
1:35652831-35653691	SFPQ (hsa-SFPQ_0037)
20:30732940-30733153	TM9SF4

Table 9 (Cont'd)

2:136505837-136530117	UBXN4 (hsa-UBXN4_0022)
3:47663697-47680270	SMARCC1 (hsa-SMARCC1_0045)
3:49397297-49397427	RHOA
6:18236683-18258636	DEK (hsa-DEK_0003)
7:102962379-102963241	DNAJC2 (hsa-DNAJC2_0003)
7:97946549-97946668	BAIAP2L1
8:61763591-61763663	CHD7
8:86050270-86050710	LRRCC1
MT:1690-1894	MT-RNR2

There are 29 exonic circRNAs which were detected in both HUH-7 cell line treated with PI3K-β and DMSO as seen in Table 10.

Table 10 Common CircNAs in HUH-7 cell line treated with both PI3K-β and DMSO

CircRNA ID	Gene Name / Annotation	Circular Expression Changes
10:95353711-95360223	RBP4	decrease

Table 10 (Cont'd)

12:103306569-103306676	PAH	decrease
12:103306569-103311003	PAH	increase
12:57108191-57108418	NACA	increase
15:53815415-53815519	WDR72	decrease
15:64448066-64448116	SNX22	decrease
17:57886157-57889147	VMP1 (hsa-VMP1_0017)	increase
17:79539043-79589304	NPLOC4	increase
17:79801680-79801750	RP11-498C9.2	decrease
18:56008910-56024484	NEDD4L	decrease
1:117944808-117963271	MAN1A2 (hsa-MAN1A2_0003)	increase
1:1586823-1650894	CDK11A|CDK11B (hsa-RP1-283E3_0030)	decrease
1:212148508-212161345	INTS7 (hsa-INTS7_0014)	decrease
1:92446200-92446938	BRDT (hsa-BRDT_0006)	increase
20:32879294-32880197	AHCY	decrease
20:57485407-57485456	GNAS	decrease

Table 10 (Cont'd)

22:42912017-42973104	RRP7A|RRP7B (hsa-RNU6-513P_0001)	decrease
2:113057426-113057606	ZC3H6 (hsa-ZC3H6_0005)	increase
3:169694734-169706147	SEC62 (hsa-SEC62_0004)	increase
4:74285971-74286038	ALB	increase
5:40832650-40834308	RPL37	decrease
7:100482834-100482991	SRRT	increase
7:100482837-100482991	SRRT	increase
7:158662546-158669382	WDR60 (hsa-WDR60_0005)	decrease
8:17503423-17503518	MTUS1	decrease
8:27462605-27462666	CLU	decrease
MT:10924-11206	MT-ND4	decrease
MT:13653-13854	MT-ND5	increase
MT:8927-9189	MT-ATP6	decrease
X:133087077-133087238	GPC3	decrease

There are 9 upregulated genes such as SEC62, PAH, GPC3, NEDD4L, ALB, AHCY, NACA, MT-ND5 and MT-ND4 in RNA sequences of HUH-7 cell lines treated with PI3K-β. In miRTargetLink, ALB, SEC62, NACA, ZC3H6, BRDT, PAH have found

as having miRNA interaction. All the genes have detected circRNAs that have emerged after the treatment or having bigger CLR results than the DMSO treatment.

The circRNA located in SEC62 gene region is one of the annotated circRNAs as known as hsa_circ_0001358 in CircBase. According to CircInteractome, it has 18 different possible and annotated miRNAs interaction such as hsa-miR-1238, hsa-miR-1248, hsa-miR-1253, hsa-miR-1279, hsa-miR-1283, hsa-miR-1296, hsa-miR-140-3p, hsa-miR-197, hsa-miR-217, hsa-miR-487a, hsa-miR-578, hsa-miR-579, hsa-miR-587, hsa-miR-587, hsa-miR-625, hsa-miR-651, hsa-miR-654-3p, hsa-miR-892b, hsa-miR-942. Although HSA_CIRCpedia_362552 and HSA_CIRCpedia_115139 located in BRDT gene in 1:92446200-92446938 and ZC3H13 gene in 13:46577274-46619651 respectively, there are not any miRNA interaction in circRNA databases like CircAtlas and CircInteractome. Gene ontology analysis of linear isoforms by using Enrichr. Like PI3K-α, the only significant GO biological process is the positive regulation of endocytosis for linear isoforms. The genes corresponding to CircRNAs that may have possible miRNA interactions have so many significant GO biological processes such as protein targeting to membrane (NACA, SEC62), catechol-containing compound biosynthetic process (PAH), erythrose 4-phosphate/phosphoenolpyruvate family amino acid metabolic process (PAH), posttranslational protein targeting to membrane, translocation (SEC62), tyrosine metabolic process (PAH), etc.

4.2.1.3 HUH-7 Cell Line treated with Sorafenib

There are 64 reliable circRNAs in HUH-7 cell line treated with Sorafenib, but only 43 of them are exonic circRNAs. There are 13 exonic circRNAs which were detected in only Sorafenib treatment that does not contain DMSO in Table 11.

Table 11 CircRNAs in HUH-7 treated with Sorafenib excluding DMSO (only Sorafenib, not containing DMSO)

CircRNA ID	Gene Name / Annotation
12:11199619-11199787	PRR4|TAS2R14 (hsa-PRR4_0005)
13:32652971-32653170	FRY (hsa-FRY_0007)
15:44843074-44843720	EIF3J (hsa-EIF3J_0001)
19:39216523-39217653	ACTN4
1:117944808-117957453	MAN1A2 (hsa-MAN1A2_0002)
22:31476436-31588688	RNF185|SMTN
2:183866681-183866771	NCKAP1 (hsa_circ_0057291)
2:64083440-64085070	UGP2 (hsa-UGP2_0001)
3:145820542-145924548	PLOD2|PLSCR4
4:103571694-103571858	MANBA
6:119611848-119628157	MAN1A1 (hsa-MAN1A1_0005)
7:90034833-90034922	CLDN12|CTB-13L3.1
9:116345997-116346050	RGS3

There are 19 exonic circRNAs which were detected in HUH-7 treated with only DMSO just before the treatment of Sorafenib in Table 12.

Table 12 CircRNAs detected in HUH-7 cell line treated with DMSO (not containing Sorafenib)

CircRNA ID	Gene Name / Annotation
10:70496711-70496805	CCAR1
15:45250571-45315547	C15orf43|SORD
15:64448066-64448116	SNX22
17:57886157-57889147	VMP1 (hsa-VMP1_0017)
1:117944808-117948267	MAN1A2 (hsa-MAN1A2_0004)
1:117944808-117963271	MAN1A2 (hsa-MAN1A2_0003)
1:228286421-228286476	ARF1
1:35652831-35653691	SFPQ (hsa-SFPQ_0037)
1:92446200-92446938	BRDT (hsa-BRDT_0006)
22:42912017-42973104	RRP7A|RRP7B (hsa-RNU6-513P_0001)

Table 12 (Cont'd)

2:113057426-113057606	ZC3H6 (hsa-ZC3H6_0005)
2:136505837-136530117	UBXN4 (hsa-UBXN4_0022)
3:169694734-169706147	SEC62 (hsa-SEC62_0004)
3:47663697-47680270	SMARCC1 (hsa-SMARCC1_0045)
6:18236683-18258636	DEK (hsa-DEK_0003)
8:86050270-86050710	LRRCC1
MT:10924-11206	MT-ND4
MT:13653-13854	MT-ND5
MT:8927-9189	MT-ATP6

Moreover, there are 30 exonic circRNAs which were detected in both Sorafenib and DMSO samples in Table 13.

Table 13 Common CircRNAs detected in both Sorafenib and DMSO samples

CircRNA ID	Gene Name / Annotation	Circular Expression Changes
10:95353711-95360223	RBP4	decrease

Table 13 (Cont'd)

10:95353731-95360223	RBP4	increase
12:103232126-103232341	PAH	decrease
12:103306569-103306676	PAH	increase
12:103306569-103311003	PAH	increase
12:57108191-57108418	NACA	decrease
15:53815415-53815519	WDR72	increase
16:418328-418560	MRPL28 (hsa-MRPL28_0001)	increase
17:79539043-79589304	NPLOC4	increase
17:79801680-79801750	RP11-498C9.2	decrease
18:56008910-56024484	NEDD4L	decrease
1:155571864-155695810	DAP3|DAP3P1 (hsa-RP11-243J18_0001)	decrease
1:1586823-1650894	CDK11A|CDK11B (hsa-RP1-283E3_0030)	decrease
1:212148508-212161345	INTS7 (hsa-INTS7_0014)	decrease
20:30732940-30733153	TM9SF4	decrease
20:32879294-32880197	AHCY	decrease
20:57485407-57485456	GNAS	increase
3:49397297-49397427	RHOA	decrease

Table 13 (Cont'd)

4:74285971-74286038	ALB	increase
5:40832650-40834308	RPL37	increase
7:100482834-100482991	SRRT	increase
7:100482837-100482991	SRRT	increase
7:102962379-102963241	DNAJC2 (hsa-DNAJC2_0003)	decrease
7:158662546-158669382	WDR60 (hsa-WDR60_0005)	decrease
7:97946549-97946668	BAIAP2L1	increase
8:17503423-17503518	MTUS1	decrease
8:27462605-27462666	CLU	increase
8:61763591-61763663	CHD7	decrease
MT:1690-1894	MT-RNR2	decrease
X:133087077-133087238	GPC3	decrease

22 upregulated genes (BAIAP2L1, FRY, NEDD4L, WDR72, NPLOC4, DNAJC2, ALB, MAN1A1, EIF3J, PAH, CHD7, MANBA, MAN1A2, GPC3, MT-RNR2, INTS7, WDR60, UGP2, TM9SF4, RGS3, ACTN4, NCKAP1) were found as linear isoforms of circRNAs in RNA sequences of HUH-7 cell line treated with Sorafenib. Except to MT-RNR2, the other genes have possible miRNA interaction according to miRTargetLink. NCKAP1, MAN1A1, UGP2, FRY, MANBA, ACTN4, MAN1A2, EIF3J are the genes that have circRNAs which emerged after the treatment or increased CLR score. hsa_circ_0057291 is in NCKAP1 gene location and it has one

possible miRNA interaction (hsa-miR-630) according to CircInteractome. hsa_circ_0000117 and hsa_circ_0001020 are the annotated circRNAs located in MAN1A1 and UGP2 respectively. They have so many possible miRNA interactions in CircInteractome.

According to Enrichr, protein alpha-1,2-demannosylation, positive regulation of endocytosis, glycoprotein metabolic process, cellular protein metabolic process, carboxylic acid biosynthetic process, ubiquitin-dependent ERAD pathway, N-glycan processing, ERAD pathway and positive regulation of supramolecular fiber organization are significant GO Biological Processes for linear isoforms of circRNAs in HUH-7 cell line treated with Sorafenib. 8 genes that may interact circRNA and miRNA, also have protein alpha-1,2-demannosylation, glycoprotein metabolic process, cellular protein metabolic process with lower p value.

4.2.1.4 HUH-7 Cell Line treated with Sorafenib+PI3K-α inhibition

There are 66 reliable circRNAs which are detected in HUH-7 treated with Sorafenib+PI3K-α inhibition, but only 48 of them are exonic circRNAs. 18 exonic circRNAs were detected in only Sorafenib+PI3K-α inhibition treated HUH-7 cell line (not containing DMSO). 8 exonic circRNAs among 18 circRNAs are annotated ones as seen in Table 14.

Table 14 CircRNAs detected in HUH-7 cell line treated with Sorafenib+PI3K-α inhibition (only Sorafenib+PI3K-α inhibition, not containing DMSO)

CircRNA ID	Gene Name / Annotation
10:70496714-70496805	CCAR1
12:114392936-114393010	RBM19 (hsa_circ_30158)
14:95053745-95053816	SERPINA5

Table 14 (Cont'd)

15:44843074-44843720	EIF3J (hsa-EIF3J_0001)
17:76000251-76027923	TNRC6C (hsa-TNRC6C_0002)
19:10782034-10782211	ILF3
19:18047219-18047388	CCDC124 (hsa-CCDC124_0001)
22:30221217-30221685	ASCC2
22:31487377-31487468	SMTN
2:183866681-183866771	NCKAP1 (hsa_circ_0057291)
3:145820542-145924548	PLOD2|PLSCR4
5:142500553-142513670	ARHGAP26 (hsa-ARHGAP26_0019)
6:119611848-119628157	MAN1A1 (hsa-MAN1A1_0005)
7:90034833-90034922	CLDN12|CTB-13L3.1
8:146016756-146016880	RPL8
8:62563609-62563683	ASPH
MT:12360-12563	MT-ND5

In HUH-7 cell line treated with only DMSO that does not contain Sorafenib+PI3K-α inhibition, 19 exonic circRNAs were detected as in Table 15.

Table 15 CircRNAs in HUH-7 cell line just before the treatment (Only DMSO)

CircRNA ID	Gene Name / Annotation
12:103232126-103232341	PAH
12:57108191-57108418	NACA
16:418328-418560	MRPL28 (hsa-MRPL28_0001)
17:57886157-57889147	VMP1 (hsa-VMP1_0017)
1:228286421-228286476	ARF1
1:35652831-35653691	SFPQ (hsa-SFPQ_0037)
1:92446200-92446938	BRDT (hsa-BRDT_0006)
20:30732940-30733153	TM9SF4
2:113057426-113057606	ZC3H6 (hsa-ZC3H6_0005)
2:136505837-136530117	UBXN4 (hsa-UBXN4_0022)
3:169694734-169706147	SEC62 (hsa-SEC62_0004)

Table 15 (Cont'd)

6:18236683-18258636	DEK (hsa-DEK_0003)
7:158662546-158669382	WDR60 (hsa-WDR60_0005)
8:17503423-17503518	MTUS1
8:86050270-86050710	LRRCC1
MT:10924-11206	MT-ND4
MT:13653-13854	MT-ND5
MT:1690-1894	MT-RNR2
X:133087077-133087238	GPC3

There are 30 common exonic circRNAs detected in HUH-7 cell line treated with Sorafenib + PI3K-α inhibition and DMSO in Table 16.

Table 16 Common CircRNAs detected in Sorafenib+PI3K-α inhibition and DMSO samples

CircRNA ID	Gene Name / Annotation	Circular Expression Changes
10:70496711-70496805	CCAR1	decrease
10:95353711-95360223	RBP4	decrease
10:95353731-95360223	RBP4	decrease
12:103306569-103306676	PAH	increase

Table 16 (Cont'd)

12:103306569-103311003	PAH	decrease
15:45250571-45315547	C15orf43|SORD	decrease
15:53815415-53815519	WDR72	increase
15:64448066-64448116	SNX22	decrease
17:79539043-79589304	NPLOC4	increase
17:79801680-79801750	RP11-498C9.2	decrease
18:56008910-56024484	NEDD4L	increase
1:117944808-117948267	MAN1A2 (hsa-MAN1A2_0004)	decrease
1:117944808-117963271	MAN1A2 (hsa-MAN1A2_0003)	increase
1:155571864-155695810	DAP3|DAP3P1 (hsa-RP11-243J18_0001)	decrease
1:1586823-1650894	CDK11A|CDK11B (hsa-RP1-283E3_0030)	decrease
1:212148508-212161345	INTS7 (hsa-INTS7_0014)	increase
20:32879294-32880197	AHCY	decrease
20:57485407-57485456	GNAS	increase
22:42912017-42973104	RRP7A|RRP7B (hsa-RNU6-513P_0001)	decrease

Table 16 (Cont'd)

3:47663697-47680270	SMARCC1 (hsa-SMARCC1_0045)	decrease
3:49397297-49397427	RHOA	decrease
4:74285971-74286038	ALB	increase
5:40832650-40834308	RPL37	decrease
7:100482834-100482991	SRRT	decrease
7:100482837-100482991	SRRT	increase
7:102962379-102963241	DNAJC2 (hsa-DNAJC2_0003)	decrease
7:97946549-97946668	BAIAP2L1	decrease
8:27462605-27462666	CLU	increase
8:61763591-61763663	CHD7	increase
MT:8927-9189	MT-ATP6	decrease

24 upregulated genes corresponding to circRNAs have been found in RNA-sequences in HUH-7 cell lines treated with Sorafenib+PI3K-α. Among 24 genes, only 3 of them have not any miRNA interaction according to miRTargetLink. 17 genes of the 24 genes (PAH, GNAS, ILF3, CLU, SMTN, WDR72, ASCC2, EIF3J, ARHGAP26, MAN1A2, RPL8, NCKAP1, CCAR1, ALB, NEDD4L, SRRT, RBM19) may be affected due to circRNA-miRNA interaction. Since the expressions of circRNAs in HUH-7 cell line treated with Sorafenib+PI3K-α have higher than DMSO samples. According to CircAtlas and CircInteractome, two circRNAs have known miRNA interactions through EIF3J, has-EIF3J 0001 and has-miR-136-3p and

MAN1A2, has-MAN1A2 003 and hsa-miR-103a-2-5p/ hsa-miR-146a-3p/ hsa-miR-135a-3p.

By using Enrichr, gene ontology analysis has been done for all genes corresponding to all detected circRNAs in HUH-7 cell line treated with Sorafenib+PI3K-α. The only significant biological process is ncRNA processing. MAN1A2 is related to N-glycan processing and protein alpha-1,2-demannosylation according to Enrichr.

4.2.1.5 HUH-7 Cell Line treated with Sorafenib+PI3K- β inhibition

62 reliable circRNAs were detected in HUH-7 Cell Line treated with Sorafenib+PI3K- β inhibition, but 45 of them are exonic circRNAs. There are 16 exonic circRNAs detected in HUH-7 cell line treated with Sorafenib+PI3K- β inhibition that don't belong DMSO as seen in Table 17.

Table 17 CircRNAs detected in HUH-7 cell line in only Sorafenib+PI3K- β inhibition (not in DMSO)

CircRNA ID	Gene Name / Annotation
10:70719562-70720005	DDX21 (hsa-DDX21_0005)
10:70925798-70928288	VPS26A
11:102319543-102319599	TMEM123
13:32652971-32653170	FRY (hsa-FRY_0007)
15:44843074-44843720	EIF3J (hsa-EIF3J_0001)

Table 17 (Cont'd)

17:76000251-76027923	TNRC6C (hsa-TNRC6C_0002)
19:18047219-18047388	CCDC124 (hsa-CCDC124_0001)
19:3734364-3735611	TJP3
1:117944808-117957453	MAN1A2 (hsa-MAN1A2_0002)
22:30736684-30737728	SF3A1 (hsa-SF3A1_0010)
22:31476436-31588688	RNF185|SMTN
3:145820542-145924548	PLOD2|PLSCR4
5:142500553-142513670	ARHGAP26 (hsa-ARHGAP26_0019)
6:119611848-119628157	MAN1A1 (hsa-MAN1A1_0005)
6:18256592-18258636	DEK (hsa-DEK_0018)
7:90034833-90034922	CLDN12|CTB-13L3.1

20 exonic circRNAs were detected in HUH-7 cell line treated with only DMSO, differ from Sorafenib+PI3K- β inhibition in Table 18.

Table 18 CircRNAs detected in HUH-7 cell line treated with only DMSO (differ than Sorafenib+PI3K- β inhibition)

CircRNA ID	Gene Name / Annotation
12:103232126-103232341	PAH
12:57108191-57108418	NACA
16:418328-418560	MRPL28 (hsa-MRPL28_0001)
1:117944808-117948267	MAN1A2 (hsa-MAN1A2_0004)
1:228286421-228286476	ARF1
1:35652831-35653691	SFPQ (hsa-SFPQ_0037)
1:92446200-92446938	BRDT (hsa-BRDT_0006)
20:32879294-32880197	AHCY
2:113057426-113057606	ZC3H6 (hsa-ZC3H6_0005)
2:136505837-136530117	UBXN4 (hsa-UBXN4_0022)
3:169694734-169706147	SEC62 (hsa-SEC62_0004)
3:47663697-47680270	SMARCC1 (hsa-SMARCC1_0045)

Table 18 (Cont'd)

5:40832650-40834308	RPL37
6:18236683-18258636	DEK (hsa-DEK_0003)
7:102962379-102963241	DNAJC2 (hsa-DNAJC2_0003)
7:158662546-158669382	WDR60 (hsa-WDR60_0005)
7:97946549-97946668	BAIAP2L1
8:86050270-86050710	LRRCC1
MT:13653-13854	MT-ND5
X:133087077-133087238	GPC3

There are 29 exonic circRNAs detected in HUH-7 cell lines treated with both Sorafenib+PI3K- β and DMSO as seen in Table 19.

Table 19 Common CircRNAs detected in HUH-7 cell lines treated with both Sorafenib+PI3K- β and DMSO

CircRNA ID	Gene Name / Annotation	Circular Expression Changes
10:70496711-70496805	CCAR1	decrease
10:95353711-95360223	RBP4	decrease
10:95353731-95360223	RBP4	decrease
12:103306569-103306676	PAH	decrease

Table 19 (Cont'd)

12:103306569-103311003	PAH	increase
15:45250571-45315547	C15orf43|SORD	decrease
15:53815415-53815519	WDR72	increase
15:64448066-64448116	SNX22	increase
17:57886157-57889147	VMP1 (hsa-VMP1_0017)	decrease
17:79539043-79589304	NPLOC4	decrease
17:79801680-79801750	RP11-498C9.2	increase
18:56008910-56024484	NEDD4L	decrease
1:117944808-117963271	MAN1A2 (hsa-MAN1A2_0003)	increase
1:155571864-155695810	DAP3|DAP3P1 (hsa-RP11-243J18_0001)	decrease
1:1586823-1650894	CDK11A|CDK11B (hsa-RP1-283E3_0030)	decrease
1:212148508-212161345	INTS7 (hsa-INTS7_0014)	decrease
20:30732940-30733153	TM9SF4	decrease
20:57485407-57485456	GNAS	decrease

Table 19 (Cont'd)

22:42912017-42973104	RRP7A|RRP7B (hsa-RNU6-513P_0001)	decrease
3:49397297-49397427	RHOA	decrease
4:74285971-74286038	ALB	increase
7:100482834-100482991	SRRT	decrease
7:100482837-100482991	SRRT	increase
8:17503423-17503518	MTUS1	decrease
8:27462605-27462666	CLU	increase
8:61763591-61763663	CHD7	decrease
MT:10924-11206	MT-ND4	increase
MT:1690-1894	MT-RNR2	decrease
MT:8927-9189	MT-ATP6	decrease

There are 16 upregulated genes corresponding to detected circRNAs in HUH-7 cell line treated with Sorafenib and PI3K-β. All the genes have possible miRNA interaction according to miRTargetLink. There are 7 different circRNAs that emerged or increased CLR score after the treatment. There are 2 known miRNA and circRNAs interactions such as hsa-let-7e-3p and hsa-MAN1A1_0005, and hsa-miR-210-3p and hsa-FRY_0007. According to Enrichr, the only significant GO biological process is the protein alpha-1,2-demannosylation for all linear isoforms of circRNAs. MAN1A1 is involved in protein deglycosylation involved in glycoprotein catabolic process, mannose trimming involved in glycoprotein ERAD pathway and

ubiquitin-dependent glycoprotein ERAD pathway and FRY is related to plasma membrane bounded cell projection organization as reported by Enrichr.

4.2.1.6 PI3K-α and PI3K-β inhibition

There are 4 common exonic circRNAs which were found in both HUH-7 cell line treated with PI3K-α and PI3K-β inhibition in Table 20.

Table 20 Common circRNAs founded in HUH-7 treated with both PI3K-α and PI3K- β

CirRNA ID	Gene Name / Annotation
15:44843074-44843720	EIF3J (hsa-EIF3J_0001)
1:117944808-117957453	MAN1A2 (hsa-MAN1A2_0002)
22:31476436-31588688	RNF185|SMTN
MT:12360-12563	MT-ND5

4.2.1.7 Sorafenib, PI3K-α inhibition and Sorafenib+PI3K-α inhibition

4 common circRNAs was found in HUH-7 cell line treated with both sorafenib and PI3K- α as seen in the Table 21.

Table 21 Common circRNAs in HUH-7 cell line treated with both Sorafenib and PI3K- α

CircRNA ID	Gene Name / Annotation
15:44843074-44843720	EIF3J (hsa-EIF3J_0001)
19:39216523-39217653	ACTN4
1:117944808-117957453	MAN1A2 (hsa-MAN1A2_0002)
22:31476436-31588688	RNF185|SMTN

Now 5 common circRNAs was found in HUH-7 cell line treated with both sorafenib and sorafenib + PI3K- α as seen in the Table 22.

Table 22 Common circRNAs found in HUH-7 treated with sorafenib and sorafenib + PI3K- α

CircRNA ID	Gene Name / Annotation
15:44843074-44843720	EIF3J (hsa-EIF3J_0001)
2:183866681-183866771	NCKAP1 (hsa_circ_0057291)
3:145820542-145924548	PLOD2|PLSCR4
6:119611848-119628157	MAN1A1 (hsa-MAN1A1_0005)
7:90034833-90034922	CLDN12|CTB-13L3.1

Table 23 reports 1 common circRNAs found in HUH-7 cell line treated with PI3K-α, sorafenib and sorafenib + PI3K- α.

Table 23 Common circRNAs found in HUH-7 treated with PI3K- α, sorafenib and sorafenib + PI3K- α

CircRNA ID	Gene Name / Annotation
15:44843074-44843720	EIF3J

4.2.1.8 Sorafenib, PI3K-β inhibition and Sorafenib+PI3K-β inhibition

Table 24 presents 10 common circRNAs found in HUH-7 cell line treated with both PI3K- β and sorafenib.

Table 24 Common circRNAs found in HUH-7 treated with sorafenib and PI3K- β

CircRNA ID	Gene Name / Annotation
15:44843074-44843720	EIF3J (hsa-EIF3J_0001)
1:117944808-117957453	MAN1A2 (hsa-MAN1A2_0002)
22:31476436-31588688	RNF185|SMTN
3:145820542-145924548	PLOD2|PLSCR4
7:90034833-90034922	CLDN12|CTB-13L3.1

7 common circRNAs was found in HUH-7 cell line treated with both sorafenib and sorafenib+ PI3K- β as seen in the Table 25.

Table 25 Common circRNAs founded in HUH-7 treated with sorafenib and sorafenib+ PI3K-β

CircRNA ID	Gene Name / Annotation
13:32652971-32653170	FRY (hsa-FRY_0007)
15:44843074-44843720	EIF3J (hsa-EIF3J_0001)
1:117944808-117957453	MAN1A2 (hsa-MAN1A2_0002)
22:31476436-31588688	RNF185|SMTN
3:145820542-145924548	PLOD2|PLSCR4
6:119611848-119628157	MAN1A1 (hsa-MAN1A1_0005)
7:90034833-90034922	CLDN12|CTB-13L3.1

Now 7 common circRNAs was found in HUH-7 cell line treated with both PI3K-β and sorafenib+ PI3K- β as seen in the Table 26.

Table 26 Common circRNAs founded in HUH-7 treated with PI3K-β and sorafenib+ PI3K-β

CircRNA ID	Gene Name / Annotation
15:44843074-44843720	EIF3J (hsa-EIF3J_0001)
17:1303341-1303755	YWHAE (hsa-YWHAE_0019)
17:76000251-76027923	TNRC6C (hsa-TNRC6C_0002)
1:117944808-117957453	MAN1A2 (hsa-MAN1A2_0002)
22:31476436-31588688	RNF185|SMTN
3:145820542-145924548	PLOD2|PLSCR4
7:90034833-90034922	CLDN12|CTB-13L3.1

Now 6 common circRNAs was found in HUH-7 cell line treated with PI3K-β, sorafenib and sorafenib + PI3K-β as seen in the Table 27.

Table 27 Common circRNAs founded in HUH-7 treated with PI3K-β, sorafenib and sorafenib + PI3K-β

CircRNA ID	Gene Name / Annotation
15:44843074-44843720	EIF3J (hsa-EIF3J_0001)

Table 27 (Cont'd)

17:1303341-1303755	YWHAE (hsa-YWHAE_0019)
1:117944808-117957453	MAN1A2 (hsa-MAN1A2_0002)
22:31476436-31588688	RNF185|SMTN
3:145820542-145924548	PLOD2|PLSCR4
7:90034833-90034922	CLDN12|CTB-13L3.1

4.2.2 Mahlavu Cell Line with different treatment options

4.2.2.1 Mahlavu Cell Line treated with PI3K-α inhibition

There are 51 and 55 reliable circRNAs detected in Mahlavu cell line treated with PI3K-α and DMSO. Among 51 circRNAs, there are 31 exonic circRNAs and 34 exonic circRNAs in DMSO treatment. There are 13 exonic circRNAs that have been detected in only PI3K-α treatment as seen Table 28.

Table 28 CircRNAs in Mahlavu Cell Line in only PI3K-α (Not containing DMSO)

circRNA ID	Gene Name / Annotation
10:27311487-27337908	ANKRD26 (hsa-ANKRD26_0006)
11:20070533-20070589	NAV2|NAV2-AS2
15:64448255-64448329	PPIB|SNX22

Table 28 (Cont'd)

16:1859755-1859834	HAGH (hsa_circ_30332)
16:70301647-70301708	AARS
17:79801680-79801750	RP11-498C9.2
19:3978111-3979347	EEF2
19:3980004-3980540	EEF2 (hsa-EEF2_0033)
1:222837407-222837490	MIA3
4:128995615-128999117	LARP1B (hsa-LARP1B_0015)
5:40832650-40834308	RPL37
6:4119420-4119509	ECI2
MT:10924-11206	MT-ND4

Table 29 reports that 15 exonic circRNAs detected in Mahlavu cell line treated with only DMSO that does not contain in PI3K-α.

Table 29 CircRNAs in Mahlavu Cell Line in only PI3K-α (Not containing DMSO)

circRNA ID	Gene Name / Annotation
11:117075229-117075388	PCSK7|TAGLN
11:1774259-1774356	CTSD
12:114392936-114393016	RBM19 (hsa_circ_30158)

Table 29 (Cont'd)

16:66967902-66968159	FAM96B
1:160183027-160183383	PEA15 (hsa-PEA15_0003)
3:196214270-196215554	RNF168 (hsa-RNF168_0001)
3:50098959-50099419	RBM6
5:131944371-131944901	RAD50 (hsa-RAD50_0044)
5:137756590-137759837	KDM3B (hsa-KDM3B_0045)
5:176764401-176764731	LMAN2
6:29797195-29856519	HLA-G|HLA-H
7:5037448-5037662	RNF216P1
8:61763591-61763663	CHD7
9:130941184-130941267	CIZ1
9:36215326-36216366	GNE

There are 18 common exonic circRNAs detected in both PI3K-α and DMSO. In Table 30 shows also that change in circular expressions after the PI3K-α treatment.

Table 30 Common CircRNAs in Mahlavu Cell Line treated with both PI3K-α and DMSO

CircRNA ID	Gene Name / Annotation	Circular Expression Changes
10:70496711-70496805	CCAR1	increase
11:64535043-64535118	SF1	decrease
14:21860666-21860758	CHD8	increase
16:30078555-30078611	ALDOA	decrease
1:1586823-1650894	CDK11A (hsa-RP1-283E3_0030)	decrease
1:6158912-6158974	KCNAB2	decrease
1:86048426-86048494	CYR61	decrease
20:30732940-30733153	TM9SF4	decrease
20:57485407-57485456	GNAS	increase
2:113057426-113057606	ZC3H6 (hsa-ZC3H6_0005)	increase
2:201337724-201337775	SPATS2L	increase
2:36776331-36776392	CRIM1	increase
3:48509707-48509793	SHISA5	increase
3:49159275-49159414	LAMB2	increase
3:49397297-49397427	RHOA	decrease

Table 30 (Cont'd)

5:149754517-149758971	TCOF1 (hsa-TCOF1_0004)	decrease
8:62563609-62563683	ASPH	increase
MT:13653-13854	MT-ND5	decrease

There are 4 upregulated genes in Mahlavu cell line treated with PI3K-α inhibition. Only 2 of them have circular isoforms whose expression increased after the treatment. The only miRNA interaction is the ANKRD26 through hsa-ANKRD26_0006 and hsa-miR-24-1-5p by using circAtlas database. proton-transporting V-type ATPase complex assembly, positive regulation of protein localization to cell surface, positive regulation of protein exit from endoplasmic reticulum and regulation of protein exit from endoplasmic reticulum are the most significant biological processes of 4 upregulated genes by using Enrichr. As seen in Figure 10, differentially expressed circRNAs and their linear isoforms are located in soma interaction pathway.

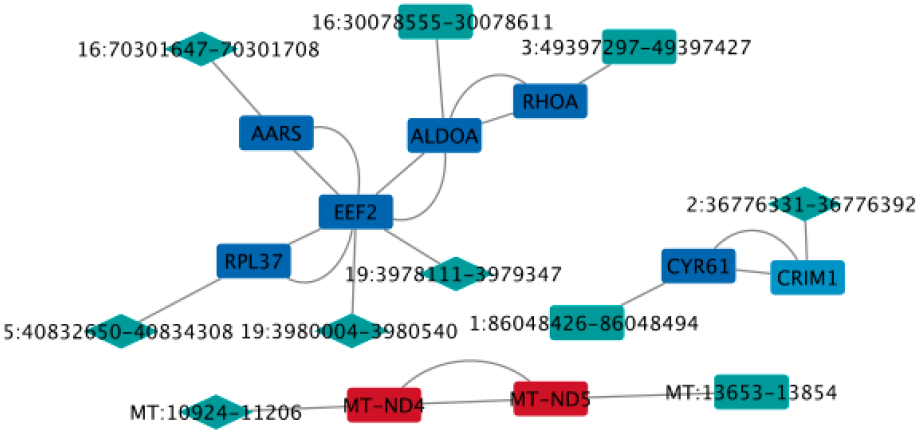

Figure 10 Differentially Expressed CircRNAs and Their Linear Isoforms in Mahlavu Cell Line Treated with PI3K- α Inhibition

4.2.2.2 Mahlavu Cell Line treated with PI3K- β inhibition

While 65 reliable circRNAs were detected in Mahlavu cell line treated with PI3K-β, there are 43 exonic circRNAs. Also, as seen Table 31, there are 10 exonic circRNAs that have emerged after the PI3K-β treatment.

Table 31 CircRNAs detected in Mahlavu Cell Line Treated With only PI3K-β (not containing DMSO)

CircRNA ID	Gene Name / Annotation
11:47489841-47489931	CELF1
14:69341426-69341587	ACTN1
19:10973846-11152236	C19orf38|SMARCA4
19:15225631-15225703	SYDE1
19:39216523-39217653	ACTN4
20:32879294-32880197	AHCY
2:176794699-176794811	KIAA1715
5:40832650-40834308	RPL37
2:176794699-176794811	KIAA1715
5:40832650-40834308	RPL37
6:47220992-47254331	TNFRSF21 (hsa-TNFRSF21_0002)
X:122753252-122756711	THOC2

20 exonic circRNAs were detected in DMSO treated Mahlavu cell line, this means they have disappeared after the PI3K-β inhibition in Table 32.

Table 32 CircRNAs in Mahlavu cell line treated with only DMSO. (not containing PI3K-β)

circRNA ID	Gene Name / Annotation
10:70496711-70496805	CCAR1
11:1774259-1774356	CTSD
12:114392936-114393010	RBM19 (hsa_circ_30158)
14:21860666-21860758	CHD8
1:160183027-160183383	PEA15 (hsa-PEA15_0003)
20:30732940-30733153	TM9SF4
2:113057426-113057606	ZC3H6 (hsa-ZC3H6_0005)
1:6158912-6158974	KCNAB2
2:201337724-201337775	SPATS2L
2:36776331-36776392	CRIM1
3:196214270-196215554	RNF168 (hsa-RNF168_0001)
3:49159275-49159414	LAMB2

Table 32 (Cont'd)

5:131944371-131944901	RAD50 (hsa-RAD50_0044)
5:137756590-137759837	KDM3B (hsa-KDM3B_0045)
5:176764401-176764731	LMAN2
6:29797195-29856519	HLA-G|HLA-H
8:61763591-61763663	CHD7
MT:13653-13854	MT-ND5
MT:13653-13854	MT-ND5

There are 14 exonic circRNAs detected in Mahlavu cell line treated with both PI3K-β and DMSO in Table 33.

Table 33 Common CircRNAs in Mahlavu Cell Line treated with both PI3K-β and DMSO

CircRNA ID	Gene Name / Annotation	Circular Expression Changes
11:117075229-117075388	PCSK7|TAGLN	decrease
11:64535043-64535118	SF1	decrease
16:30078555-30078611	ALDOA	decrease
16:66967902-66968159	FAM96B	increase

Table 33 (Cont'd)

1:1586823-1650894	CDK11A (hsa-RP1-283E3_0030)	decrease
1:86048426-86048494	CYR61	increase
20:57485407-57485456	GNAS	increase
3:48509707-48509793	SHISA5	increase
3:49397297-49397427	RHOA	decrease
3:50098959-50099419	RBM6	increase
5:149754517-149758971	TCOF1 (hsa-TCOF1_0004)	increase
8:62563609-62563683	ASPH	decrease
9:130941184-130941267	CIZ1	increase
9:36215326-36216366	GNE	decrease

14 upregulated genes were found Mahlavu cell line treated with PI3K-β. Among 14 genes, 4 of them have circular isoforms that emerged or increased expression after the treatment. Only has-TNFRSF21_002 has known miRNA interaction between TNFRSF21 and hsa-miR-30b-3p in CircAtlas. TNFRSF21 is TNF receptor superfamily member, and it promotes the apoptosis. Moreover, the rest has no miRNA interaction according to miRTargetLink. When Enrichr gene ontology analysis has been done by considering all the upregulated genes, they involve such biological processes; negative regulation of cell-substrate adhesion, regulation of cellular component movement, positive regulation of NIK/NF-kappaB signaling, regulation of sodium:proton antiporter activity, negative regulation of cell size, lymphocyte apoptotic process.

4.2.2.3 Mahlavu Cell Line treated with Sorafenib

52 reliable circRNAs were detected in Mahlavu cell line treated with Sorafenib and 34 of them are exonic circRNAs. Moreover, 14 exonic circRNAs were found in only after Sorafenib treatment that does not contain DMSO in Table 34.

Table 34 CircRNAs in Mahlavu treated with Sorafenib excluding DMSO

CircRNA ID	Gene Name / Annotation
10:112360808-112360888	SMC3
12:57108191-57108418	NACA
12:6688064-6690252	CHD4 (hsa-CHD4_0021)
16:70301647-70301708	AARS
17:79801680-79801750	RP11-498C9.2
19:10973846-11152236	C19orf38|SMARCA4
19:3978111-3979347	EEF2
1:117944808-117984947	MAN1A2 (hsa-MAN1A2_0008)
20:32879294-32880197	AHCY
2:136527339-136530117	UBXN4 (hsa-UBXN4_0001)
4:54308820-54319107	FIP1L1
5:40832650-40834308	RPL37
6:31380102-31474914	MICA|MICB

There are 14 circRNAs detected in Mahlavu cell line just before the Sorafenib treatment in Table 35.

Table 35 CircRNAs detected in Mahlavu cell line treated with DMSO (Not containing Sorafenib)

CircRNA ID	Gene Name / Annotation
14:21860666-21860758	CHD8
16:66967902-66968159	FAM96B
1:6158912-6158974	KCNAB2
2:113057426-113057606	ZC3H6 (hsa-ZC3H6_0005)
2:201337724-201337775	SPATS2L
2:36776331-36776392	CRIM1
3:196214270-196215554	RNF168 (hsa-RNF168_0001)
3:50098959-50099419	RBM6
5:131944371-131944901	RAD50 (hsa-RAD50_0044)
5:137756590-137759837	KDM3B (hsa-KDM3B_0045)
7:5037448-5037662	RNF216P1
8:61763591-61763663	CHD7
9:36215326-36216366	GNE (hsa-GNE_0013)

Table 35 (Cont'd)

MT:13653-13854	MT-ND5

Furthermore, there are 19 common circRNAs detected in Mahlavu cell line treated with both Sorafenib and DMSO in Table 36.

Table 36 Common CircRNAs detected in Mahlavu cell line treated with both Sorafenib and DMSO

CircRNA ID	Gene Name / Annotation	Circular Expression Changes
10:70496711-70496805	CCAR1	increase
11:117075229-117075388	PCSK7|TAGLN	Decrease
11:1774259-1774356	CTSD	increase
11:64535043-64535118	SF1	decrease
12:114392936-114393010	RBM19 (hsa_circ_30158)	increase
16:30078555-30078611	ALDOA	decrease
1:1586823-1650894	CDK11A (hsa-RP1-283E3_0030)	decrease
1:160183027-160183383	PEA15 (hsa-PEA15_0003)	increase
1:86048426-86048494	CYR61	decrease
20:30732940-30733153	TM9SF4	decrease
20:57485407-57485456	GNAS	increase

Table 36 (Cont'd)

3:48509707-48509793	SHISA5	increase
3:49159275-49159414	LAMB2	increase
3:49397297-49397427	RHOA	decrease
5:149754517-149758971	TCOF1 (hsa-TCOF1_0004)	decrease
5:176764401-176764731	LMAN2	decrease
6:29797195-29856519	HLA-G|HLA-H	decrease
8:62563609-62563683	ASPH	decrease
9:130941184-130941267	CIZ1	decrease

While there are 18 upregulated genes detected in Mahlavu treated with Sorafenib, 10 of 18 upregulated genes have circular isoforms whose expression increased after the treatment. Among 10 genes, 2 of them have possible miRNA interactions such as hsa-PEA15_0003- PEA15-hsa-miR-1225-5p and hsa-UBXN4_0001- UBXN4- hsa-miR-18a-3p. 7 (AARS, GNAS, SMC3, AHCY, PEA15, CTSD, NACA) of 8 upregulated genes have possible miRNA interaction according to miRTargetLink. Mitotic spindle assembly, neutrophil degranulation and neutrophil activation involved in immune response are the significant GO biological processes for 18 upregulated genes by using Enrichr.

4.2.2.4 Mahlavu Cell Line treated with Sorafenib+PI3K-α inhibition

There are 45 reliable circRNAs detected in Mahlavu cell line treated with Sorafenib+PI3K-α inhibition. Among 45 reliable circRNAs, 26 circRNAs are exonic circRNAs. Table 37 shows that 9 circRNAs were detected in only Sorafenib+PI3K-α that does not contain in DMSO sample.

Table 37 CircRNAs detected in Mahlavu Cell Line treated with only Sorafenib+PI3K-α

CircRNA ID	Gene Name / Annotation
10:27311487-27337908	ANKRD26 (hsa-ANKRD26_0006)
14:69341426-69341587	ACTN1
17:40985505-40985702	PSME3
19:39216523-39217653	ACTN4
1:117944808-117963271	MAN1A2 (hsa-MAN1A2_0003)
20:32879294-32880197	AHCY
5:177637636-177637878	HNRNPAB (hsa-HNRNPAB_0004)
6:4119420-4119509	ECI2
8:27462605-27462666	CLU

In Table 38, there are 16 circRNAs found in DMSO samples just before Sorafenib+PI3K-α treatment.

Table 38 CircRNAs in Mahlavu Cell line just before the treatment (Only DMSO)

CircRNA ID	Gene Name / Annotation
11:117075229-117075388	PCSK7|TAGLN
12:114392936-114393010	RBM19 (hsa_circ_30158)
16:30078555-30078611	ALDOA
1:160183027-160183383	PEA15 (hsa-PEA15_0003)
1:6158912-6158974	KCNAB2
3:49159275-49159414	LAMB2
5:131944371-131944901	RAD50 (hsa-RAD50_0044)
5:137756590-137759837	KDM3B (hsa-KDM3B_0045)
5:149754517-149758971	TCOF1 (hsa-TCOF1_0004)
5:176764401-176764731	LMAN2
6:29797195-29856519	HLA-G|HLA-H
7:5037448-5037662	RNF216P1
8:61763591-61763663	CHD7
8:62563609-62563683	ASPH

Table 38 (Cont'd)

9:36215326-36216366	GNE (hsa-GNE_0013)
MT:13653-13854	MT-ND5

17 common circRNAs were detected in Mahlavu cell line treated with both Sorafenib+PI3K-α and DMSO in Table 39.

Table 39 Common circRNAs detected in both Sorafenib+PI3K-α and DMSO samples

CircRNA ID	Gene Name / Annotation	Circular Expression Changes
10:70496711-70496805	CCAR1	increase
11:1774259-1774356	CTSD	decrease
11:64535043-64535118	SF1	decrease
14:21860666-21860758	CHD8	Increase
16:66967902-66968159	FAM96B	increase
1:1586823-1650894	CDK11A (hsa-RP1-283E3_0030)	decrease
1:86048426-86048494	CYR61	increase
20:30732940-30733153	TM9SF4	increase
20:57485407-57485456	GNAS	decrease

Table 39 (Cont'd)

2:113057426-113057606	ZC3H6 (hsa-ZC3H6_0005)	decrease
2:201337724-201337775	SPATS2L	increase
2:36776331-36776392	CRIM1	increase
3:196214270-196215554	RNF168 (hsa-RNF168_0001)	increase
3:48509707-48509793	SHISA5	increase
3:49397297-49397427	RHOA	increase
3:50098959-50099419	RBM6	increase
9:130941184-130941267	CIZ1	increase

There are 13 upregulated genes in Mahlavu cell line treated with Sorafenib+PI3K-α. Among all genes, 8 genes (RHOA, FAM96B, SPATS2L, CHD8, ACTN4, TM9SF4, HNRNPAB and RNF168) have circRNAs that emerged or increased expression of circular isoforms. Has-RNF168_001 and has-HNRNPAB_004 have known miRNA interaction with hsa-miR-31-5p and hsa-miR-124-5p respectively. CCR3 signaling in eosinophils, Plasma membrane estrogen receptor signaling, Platelet activation, signaling and aggregation, Hemostasis pathway and Adherens junction cell adhesion are the significant pathways of 13 upregulated genes according to Enrichr. Although 6 genes corresponding to circRNAs have not known miRNA interaction in circRNA databases, genes including FAM96B, SPATS2L, CHD8, ACTN4 and TM9SF4 have possible miRNA interaction according to miRTargetLink.

4.2.2.5 Mahlavu Cell Line treated with Sorafenib+PI3K-β inhibition

65 reliable CircRNAs were detected in Mahlavu cell line treated with Sorafenib+PI3K-β. 42 of them are exonic circRNAs. 20 exonic circRNAs were found in only Sorafenib+PI3K-β inhibition as shown in Table 40.

Table 40 CircRNAs detected in Mahlavu cell line treated with Sorafenib+PI3K-β (not containing DMSO)

CircRNA ID	Gene Name / Annotation
10:27311487-27337908	ANKRD26 (hsa-ANKRD26_0006)
12:57108191-57108418	NACA
15:64448255-64448329	PPIB|SNX22
15:93543742-93558139	CHD2 (hsa-CHD2_0085)
16:70301647-70301708	AARS
16:77229014-77229521	MON1B
17:79801680-79801750	RP11-498C9.2
19:39216523-39217653	ACTN4
1:117944808-117963271	MAN1A2 (hsa-MAN1A2_0003)
20:34301019-34301145	RBM39
2:181925392-181925496	UBE2E3
4:128995615-128999117	LARP1B (hsa-LARP1B_0015)

Table 40 (Cont'd)

5:31526871-31526956	DROSHA
5:40832650-40834308	RPL37
6:31759994-31760078	VARS
6:47220992-47254331	TNFRSF21 (hsa-TNFRSF21_0002)
7:149992341-149992436	ACTR3C
8:27462605-27462666	CLU
X:48759233-48759319	PQBP1
4:128807782-128807878	PLK4

There are 22 common circRNAs detected in Mahlavu cell line treated with both Sorafenib+PI3K-β and DMSO in Table 41.

Table 41 Common CircRNAs detected in both Sorafenib+PI3K-β and DMSO samples

CircRNA ID	Gene Name / Annotation	Circular Expression Changes
10:70496711-70496805	CCAR1	increase
11:117075229-117075388	PCSK7	increase
11:1774259-1774356	CTSD	decrease
11:64535043-64535118	SF1	decrease
12:114392936-114393010	RBM19 (hsa_circ_30158)	increase

Table 41 (Cont'd)

14:21860666-21860758	CHD8	increase
16:30078555-30078611	ALDOA	decrease
1:1586823-1650894	CDK11A (hsa-RP1-283E3_0030)	decrease
1:6158912-6158974	KCNAB2	decrease
1:86048426-86048494	CYR61	decrease
20:30732940-30733153	TM9SF4	decrease
20:57485407-57485456	GNAS	increase
2:113057426-113057606	ZC3H6 (hsa-ZC3H6_0005)	decrease
3:196214270-196215554	RNF168 (hsa-RNF168_0001)	increase
3:48509707-48509793	SHISA5	Increase
3:49397297-49397427	RHOA	increase
5:137756590-137759837	KDM3B (hsa-KDM3B_0045)	increase
5:176764401-176764731	LMAN2	decrease
6:29797195-29856519	HLA-G	increase
8:61763591-61763663	CHD7	decrease
8:62563609-62563683	ASPH	decrease

There are 24 upregulated genes in Mahlavu cell line treated with Sorafenib+PI3K-β, but 17 of them have circRNAs that emerged and increased the circular isoform expression after the treatment. Moreover, 6 of 17 genes corresponding to circular isoforms have known miRNA-circRNA interactions such as hsa-RNF168_0001-hsa-miR-31-5p, hsa-ANKRD26_0006-hsa-miR-24-1-5p, hsa-MAN1A2_0003-hsa-miR-103a-2-5p/ hsa-miR-146a-3p/ hsa-miR-135a-3p, hsa-LARP1B_0015-hsa-miR-212-5p/ hsa-miR-146a-3p, hsa-KDM3B_0045-hsa-miR-181a-2-3p and hsa-CHD2_0085-hsa-miR-134-3p as reports in CircAtlas. Although the rest of the genes (UBE2E3, CLU, RHOA, SHISA5, CHD8, DROSHA, GNAS) corresponding to circular isoforms have no known miRNA interaction according to circRNA databases, they have possible miRNA interaction according to miRTargetLink. After doing gene ontology analysis of 24 upregulated genes, Downregulation of MTA-3 in ER-negative breast tumors and CCR3 signaling in eosinophils are the significant pathways. However, in gene ontology analysis of genes corresponding to circRNAs that have known miRNA interactions, somatic recombination of immunoglobulin genes involved in immune response, isotype switching, negative regulation of transcription elongation from RNA polymerase II promoter and negative regulation of DNA-templated transcription, elongation are the significant biological processes.

4.2.2.6 PI3K-α and PI3K-β inhibition

There is only one circRNAs that was found in Mahlavu cell line treated with both PI3K-α and PI3K-β inhibition in Table 42.

Table 42 Common circRNAs in Mahlavu cell line treated with both PI3K-α and PI3K-β

CircRNA ID	Gene Name / Annotation
5:40832650-40834308	RPL37

4.2.2.7 Sorafenib, PI3K-α inhibition and Sorafenib+PI3K-α inhibition

5 exonic circRNAs were detected in Mahlavu Cell line treated with both Sorafenib and PI3K-α in Table 43.

Table 43 Common circRNAs in Mahlavu cell line treated with both Sorafenib and PI3K- α

CircRNA ID	Gene Name / Annotation
16:70301647-70301708	AARS
17:79801680-79801750	RP11-498C9.2
19:3978111-3979347	EEF2
5:40832650-40834308	RPL37
6:4119420-4119509	ECI2

There are 2 exonic circRNAs detected in both Sorafenib and Sorafenib+ PI3K-α samples in Table 44.

Table 44 Common circRNAs in Mahlavu cell line treated in both Sorafenib and Sorafenib+ PI3K- α

CircRNA ID	Gene Name / Annotation
20:32879294-32880197	AHCY
6:4119420-4119509	ECI2

There is only circRNA detected in Mahlavu cell line treated with PI3K-α, Sorafenib and Sorafenib+ PI3K-α. It is in 6:4119420-4119509 located on ECl2 gene.

4.2.2.8 Sorafenib, PI3K-β inhibition and Sorafenib+PI3K-β inhibition

There are 3 exonic circRNAs detected in both PI3K-β inhibition and Sorafenib samples as seen Table 45.

Table 45 CircRNAs detected in both PI3K-β inhibition and Sorafenib samples

CircRNAs ID	Gene name / Annotation
19:10973846-11152236	C19orf38|SMARCA4
20:32879294-32880197	AHCY
5:40832650-40834308	RPL37

4 exonic CircRNAs were detected in both Sorafenib and Sorafenib+ PI3K-β samples in Table 46.

Table 46 CircRNAs detected in both Sorafenib and Sorafenib+ PI3K-β samples

CircRNA ID	Gene Name / Annotation
12:57108191-57108418	NACA
16:70301647-70301708	AARS
17:79801680-79801750	RP11-498C9.2
5:40832650-40834308	RPL37

Only 1 CircRNA was detected in Sorafenib, PI3K-β and Sorafenib+ PI3K-β samples. It is 5:40832650|40834308 located on RPL37 gene.

4.2.3 HUH-7 and Mahlavu Cell Lines with same treatment options

There are 12 reliable circRNAs detected in both DSMO samples of HUH-7 and Mahlavu cell lines. Among all 12 circRNAs, there are 8 exonic circRNAs as seen in Table 47.

Table 47 Common CircRNAs in both HUH-7 and Mahlavu Cell line

CircRNA ID	Gene Name / Annotation
10:70496711-70496805	CCAR1
1:1586823-1650894	CDK11A (hsa-RP1-283E3_0030)
20:30732940-30733153	TM9SF4
20:57485407-57485456	GNAS
2:113057426-113057606	ZC3H6 (hsa-ZC3H6_0005)
3:49397297-49397427	RHOA
8:61763591-61763663	CHD7
MT:13653-13854	MT-ND5

CHAPTER 5

CONCLUSION

6.1 Summary

In this study, we aimed to develop a pipeline to detect circRNAs from regular RNA-seq data using unmapped back-splice junction information and to analyze their transcript level differential expression. We selected 12 different RNA sequences of HUH-7 and Mahlavu cell lines applied with different treatment options including sorafenib, PI3K-α, PI3K- β inhibitor and their combinations as sorafenib and PI3K-α inhibitor and sorafenib and PI3K- β inhibitor. CirComPara was selected as detection and quantification method due to availability of working at the same time by using different detection methods. Due to some computational challenges, detection of circRNAs may have too much false positives. In order to remove false positives, 5 different detection methods including CIRI, CircExplorer2, DCC, findcirc, circRNAfinder, two different filtering options and circular/linear isoforms ratio were applied. Using mapped reads and unmapped reads of regular RNA-seq data, linear gene and circRNAs' expressions were generated respectively. These filtering options was to have more than two reads in sequences and to be detected at least two different detection methods. These filtering options and circular/linear isoforms ratio were mandatory approaches for our data set. Because our RNA sequences were not prepared especially for the detection, they are total RNA sequences. Among all detected circRNAs, only exonic circRNAs have been considered to investigate miRNA-mRNA interaction through parental genes into sorafenib, PI3K-α, PI3K- β inhibitor pathways. Moreover, Gene Set Enrichment Analysis has been done for samples using differential expression of linear isoforms.

6.2 Discussion

Total number of detected circRNAs were 11245 and 12783 in the HUH-7 and Mahlavu cell lines just before applying any filtering options and not considering circular/linear isoforms ratio. When we selected the circRNAs which have more than two reads in sequences and detectable with at least two methods, 136 in HUH-7 and 122 in Mahlavu cell lines circRNAs have remained as reliable circRNAs. There are 97 reliable exonic circRNAs overlapping 97 genes in Huh-7 cell lines, and 79 reliable circRNAs overlapping 84 genes in Mahlavu cell lines.

Also we did not use any positive control data in order to check whether this pipeline can find the annotated circRNAs from synthetic RNA sequences. This can be done for the future work. Because this is an important step to verify all the filtering steps.

In purification steps of library preparation, poly(A) enrichment steps or treatment of RNase R are the best methods to have minimum number of false positive reads for the detection of circRNAs. [42] However RNA-sequences that were used in this work are total RNA sequences. The study conducted by Hansen et al. (2015) in order to compare different algorithms; circRNA_finder, find_circ, CIRCexplorer, CIRI and Mapsplice by using two different treatment options (with RNase R and without RNase R) for RNA sequences reported that CIRCexplorer and Mapsplice produced the most reliable circRNAs from RNA sequences without RNase R treatment. [79] CIRCexplorer2 and DCC were two detection methods with the lowest false-positive rate for both HUH-7 and Mahlavu cell line, this verifies the study conducted by Hansen et al. (2015). Although find_circ and circRNA_finder could detect the most reliable circRNAs for both cell lines, they have lower accuracy rates compared to others. Another study had been conducted to compare 11 different detection methods including CIRCexplorer, circRNA_finder, find_circ, MapSplice, NCLScan; PTESFinder, CIRI, UROBUS, Segemehl, KNIFE, DCC by using real and simulated

datasets and they also reported that find_circ is one of the detection methods with the worst performance. [80]

In order to investigate the role of exonic circRNAs in gene regulation by sponging miRNAs, all annotated circRNAs were analyzed by using circRNA databases (circ2traits, circad, CSCD, CircPedia, circinteractome, circ2disease, circRNAdb, circAtlas). However, any relation between circRNAs and miRNAs could not be found in sorafenib, PI3K-α, PI3K- β pathways through their parental gene by using KEGG and GeneCards. When we compared to HUH-7 and Mahlavu cell lines treated with same therapeutic agents, we could not find common circRNAs. However, their samples treated with only DMSO have 8 common exonic circRNAs. One of the reasons why any annotated circRNAs related to HCC [11, 17, 67, 68, 70, 71, 72,73, 75, 77] could not found in both HUH-7 and Mahlavu cell lines may be due to use of so many filters such as being detectable at least two detection methods, having more than two reads and high circ_score.

6.3 Future Work

Despite the fact that any strong relations between the cell lines and their treatment options through miRNA-circRNA-mRNA, circRNAs have other interactions with long non-coding RNAs and RNA binding proteins and functions. Although the full mechanism is still not clear, some intronic and exonic-intronic circRNAs can increase the transcription of parental gens interaction with U1 small nuclear ribonucleoprotein particle. [51] So, not only exonic circRNAs, but also intronic and exonic-intronic circRNAs should consider for the future work. Moreover, synthetic RNA sequences as positive control data should be involved in this work to verify the accuracy of Circompara.

REFERENCES

1 Haas-Kogan, D., & Tinkle, C. (2012). Hepatocellular carcinoma: Natural history, current management, and emerging tools. Biologics: Targets and Therapy, 207. doi:10.2147/btt.s23907

2 Llovet, J., Zucman-Rossi, J., Pikarsky, E. et al. Hepatocellular carcinoma. Nat Rev Dis Primers 2, 16018 (2016) doi:10.1038/nrdp.2016.18

3 Désert R, Nieto N, Musso O. Dimensions of hepatocellular carcinoma phenotypic diversity. World J Gastroenterol 2018; 24(40): 4536-4547

4 Nault JC, Villanueva A. Intratumor molecular and phenotypic diversity in hepatocellular carcinoma. Clin Cancer Res 2015;21: 1786-1788

5 Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

6 Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A FLEXIBLE trimmer FOR Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

7 Kim, D., Paggi, J.M., Park, C. *et al.* Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol***37,** 907–915 (2019). https://doi.org/10.1038/s41587-019-0201-4

8 Xu, J., Ji, L., Liang, Y., Wan, Z., Zheng, W., Song, X., Gorshkov, K., Sun, Q., Lin, H., Zheng, X., Chen, J., Jin, R.-an, Liang, X., & Cai, X. (2020). Circrna-sore mediates sorafenib resistance in hepatocellular carcinoma by stabilizing ybx1. Signal Transduction and Targeted Therapy, 5(1). https://doi.org/10.1038/s41392-020-00375-5

9 Luo, Y., Fu, Y., Huang, R., Gao, M., Liu, F., Gui, R., & Nie, X. (2019). CircRNA_101505 sensitizes hepatocellular carcinoma cells to CISPLATIN BY sponging miR-103 and Promotes oxidored-nitro domain-containing protein 1 expression. Cell Death Discovery, 5(1). https://doi.org/10.1038/s41420-019-0202-6

10 Chakrabarti, J., & Sanga, M. (2018). Circular RNA: A Novel Noncoding RNA in Cancer. In Cancer and Noncoding Rnas: Volume 1 (pp. 187-221). London: Academic Press.

11 Xu L, Zhang M, Zheng X et al (2017) The circular RNA ciRS-7 (Cdr1as) acts as a risk factor of hepatic microvascular invasion in hepatocellular carcinoma. J Cancer Res Clin Oncol 143(1):1–11

12 López-Jiménez, E., Rojas, A. M., Andrés-León, E., & Aghaee-Bakhtiari, S. H. (2018). Bioinformatics for Circular RNAs. In Circular RNAs: Biogenesis and Functions (pp. 17-41). Singapore: Springer Verlag, Singapore. doi:10.1007/978-981-13-1426-1

13 Awasthi, R., Singh, A. K., Mishra, G., Maurya, A., Chellappan, D. K., Gupta, G., . . . Dua, K. (2018). An Overview of Circular RNAs. In Circular RNAs: Biogenesis and Functions (pp. 3-17). Singapore: Springer Verlag, Singapore. doi:10.1007/978-981-13-1426-1

14 Shang, Q., Yang, Z., Jia, R., & Ge, S. (2019). The novel roles of circRNAs in human cancer. Molecular Cancer, 18(1). doi:10.1186/s12943-018-0934-6

15 Chekulaeva, M., & Rajewsky, N. (2018). Roles of Long Noncoding RNAs and Circular RNAs in Translation. Cold Spring Harbor Perspectives in Biology, 11(6). doi:10.1101/cshperspect.a032680

16 Feng, Z., Meng, S., Zhou, H., Xu, Z., Tang, Y., Li, P., . . . Wu, M. (2019). Functions and Potential Applications of Circular RNAs in Cancer Stem Cells. Frontiers in Oncology, 9. doi:10.3389/fonc.2019.00500

17 Bach, D., Lee, S. K., & Sood, A. K. (2019). Circular RNAs in Cancer. Molecular Therapy - Nucleic Acids, 16, 118-129. doi:10.1016/j.omtn.2019.02.005

18 Zhao, X., Cai, Y., & Xu, J. (2019). Circular RNAs: Biogenesis, Mechanism, and Function in Human Cancers. International Journal of Molecular Sciences, 3926(20), 16th ser.

19 Xia, L., Song, M., Sun, M., Wang, F., & Yang, C. (2018). Circular RNAs as Biomarkers for Cancer. In Circular RNAs: Biogenesis and Functions (pp. 171-191). Singapore: Springer Verlag, Singapore.

20 Yu, T., Wang, Y., Fan, Y., Fang, N., Wang, T., Xu, T., & Shu, Y. (2019). CircRNAs in cancer metabolism: A review. Journal of Hematology & Oncology, 12(1). doi:10.1186/s13045-019-0776-8

21 Sapisochin G, Bruix J. Liver transplantation for hepatocellular carcinoma: outcomes and novel surgical approaches. Nat Rev Gastroenterol Hepatol 2017; 14: 203-217

22 Dimri, M., & Satyanarayana, A. (2020). Molecular signaling pathways and therapeutic targets in hepatocellular carcinoma. Cancers, 12(2), 491. https://doi.org/10.3390/cancers12020491

23 Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. Cell, 100(1), 57–70. https://doi.org/10.1016/s0092-8674(00)81683-9

24 Sorafenib (BAY 43-9006) inhibits tumor growth and vascularization and induces tumor apoptosis and hypoxia in RCC xenograft models. Chang YS, Adnane J, Trail PA, Levy J, Henderson A, Xue D, Bortolon E, Ichetovkin M, Chen C, McNabola A, Wilkie D, Carter CA, Taylor IC, Lynch M, Wilhelm S Cancer Chemother Pharmacol. 2007 Apr; 59(5):561-74.

25 Llovet J.M., Ricci S., Mazzaferro V., Hilgard P., Gane E., Blanc J.-F., De Oliveira A.C., Santoro A., Raoul J.-L., Forner A. Sorafenib in advanced

hepatocellular carcinoma. *N. Engl. J. Med.* 2008;359:378–390. doi: 10.1056/NEJMoa0708857.

26 Ren, S., Xin, Z., Xu, Y., Xu, J., & Wang, G. (2017). Construction and analysis of circular RNA molecular regulatory networks in liver cancer. Cell Cycle, 16(22), 2204-2211. doi:10.1080/15384101.2017.1346754

27 Guo, W., Zhang, J., Zhang, D., Cao, S., Li, G., Zhang, S., . . . Ye, H. (2017). Polymorphisms and expression pattern of circular RNA circ-ITCH contributes to the carcinogenesis of hepatocellular carcinoma. Oncotarget, 8(29), 48169-48177. doi:10.18632/oncotarget.18327

28 Song, M., Xia, L., Sun, M., Yang, C., & Wang, F. (2018). Circular RNA in Liver: Health and Diseases. In Circular RNAs: Biogenesis and Functions (pp. 245-259). Singapore: Springer Verlag, Singapore. doi:10.1007/978-981-13-1426-1

29 Glažar, P., Papavasileiou, P., & Rajewsky, N. (2014). CircBase: A database for circular RNAs. Rna, 20(11), 1666-1670. doi:10.1261/rna.043687.113

30 Xia, S., Feng, J., Lei, L., Hu, J., Xia, L., Wang, J., . . . He, C. (2016). Comprehensive characterization of tissue-specific circular RNAs in the human and mouse genomes. Briefings in Bioinformatics. doi:10.1093/bib/bbw081

31 Vo, J. N., Cieslik, M., Zhang, Y., Shukla, S., Xiao, L., Zhang, Y., . . . Chinnaiyan, A. M. (2019). The Landscape of Circular RNA in Cancer. Cell, 176(4). doi:10.1016/j.cell.2018.12.021

32 Chen, X., Han, P., Zhou, T., Guo, X., Song, X., & Li, Y. (2016). CircRNADb: A comprehensive database for human circular RNAs with protein-coding annotations. Scientific Reports, 6(1). doi:10.1038/srep34985

33 Fan, C., Lei, X., Fang, Z., Jiang, Q., & Wu, F. (2018). CircR2Disease: A manually curated database for experimentally supported circular RNAs associated with various diseases. Database, 2018. doi:10.1093/database/bay044

34 Rophina, M., Sharma, D., Poojary, M., & Scaria, V. (2020). Circad: A comprehensive manually curated resource of circular RNA associated with diseases. Database, 2020. doi:10.1093/database/baaa019

35 Zhao, Z., Wang, K., Wu, F., Wang, W., Zhang, K., Hu, H., . . . Jiang, T. (2018). CircRNA disease: A manually curated database of experimentally supported circRNA-disease associations. Cell Death & Disease, 9(5). doi:10.1038/s41419-018-0503-3

36 Ghosal, S., Das, S., Sen, R., Basak, P., & Chakrabarti, J. (2013). Circ2Traits: A comprehensive database for circular RNA potentially associated with disease and traits. Frontiers in Genetics, 4. doi:10.3389/fgene.2013.00283

37 Meng, X., Hu, D., Zhang, P., Chen, Q., & Chen, M. (2019). CircFunBase: A database for functional circular RNAs. Database, 2019. doi:10.1093/database/baz003

38 Dong, R., Ma, X., Li, G., & Yang, L. (2018). CIRCpedia v2: An Updated Database for Comprehensive Circular RNA Annotation and Expression Comparison. Genomics, Proteomics & Bioinformatics, 16(4), 226-233. doi:10.1016/j.gpb.2018.08.001

39 Ji, P., Wu, W., Chen, S., Zheng, Y., Zhou, L., Zhang, J., . . . Zhao, F. (2019). Expanded Expression Landscape and Prioritization of Circular RNAs in Mammals. Cell Reports, 26(12). doi:10.1016/j.celrep.2019.02.078

40 Xia, S., Feng, J., Chen, K., Ma, Y., Gong, J., Cai, F., . . . He, C. (2017). CSCD: A database for cancer-specific circular RNAs. Nucleic Acids Research, 46(D1). doi:10.1093/nar/gkx863

41 Dudekula, D. B., Panda, A. C., Grammatikakis, I., De, S., Abdelmohsen, K., & Gorospe, M. (2015). CircInteractome: A web tool for exploring circular RNAs and their interacting proteins and microRNAs. RNA Biology, 13(1), 34-42. doi:10.1080/15476286.2015.1128065

42 Szabo, L., & Salzman, J. (2016). Detecting circular RNAs: Bioinformatic and experimental challenges. Nature Reviews Genetics, 17(11), 679-692. doi:10.1038/nrg.2016.114

43 Li, S., & Han, L. (2019). Circular RNAs as promising biomarkers in cancer: Detection, function, and beyond. Genome Medicine, 11(1). doi:10.1186/s13073-019-0629-7

44 Zhang, X., Dong, R., Zhang, Y., Zhang, J., Luo, Z., Zhang, J., . . . Yang, L. (2016). Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. Genome Research, 26(9), 1277-1287. doi:10.1101/gr.202895.115

45 Jakobi, T., & Dieterich, C. (2019). Computational approaches for circular RNA analysis. Wiley Interdisciplinary Reviews: RNA, 10(3). doi:10.1002/wrna.1528

46 Gao, Y., Wang, J., & Zhao, F. (2015). CIRI: An efficient and unbiased algorithm for de novo circular RNA identification. Genome Biology, 16(1). doi:10.1186/s13059-014-0571-3

47 Wang, K., Singh, D., Zeng, Z., Coleman, S. J., Huang, Y., Savich, G. L., . . . Liu, J. (2010). MapSplice: Accurate mapping of RNA-seq reads for splice junction discovery. Nucleic Acids Research, 38(18). doi:10.1093/nar/gkq622

48 Szabo, L., Morey, R., Palpant, N. J., Wang, P. L., Afari, N., Jiang, C., . . . Salzman, J. (2016). Erratum to: Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development. Genome Biology, 17(1). doi:10.1186/s13059-016-1123-9

49 Sulaiman, S. A., Murad, N. A., Hanif, E. M., Abu, N., & Jamal, R. (2018). Prospective Advances in Circular RNA Investigation. In Circular RNAs: Biogenesis and Functions (pp. 357-369). Singapore: Springer Verlag, Singapore.

50 Ma, X., Wang, M., Liu, C., Dong, R., Carmichael, G. G., Chen, L., & Yang, L. (2019). A CLEAR pipeline for direct comparison of circular and linear RNA expression. doi:10.1101/668657

51 Li, Z., Huang, C., Bao, C., Chen, L., Lin, M., Wang, X., Zhong, G., Yu, B., Hu, W., Dai, L., Zhu, P., Chang, Z., Wu, Q., Zhao, Y., Jia, Y., Xu, P., Liu, H., & Shan, G. (2015). Exon-intron circular rnas regulate transcription in the nucleus. Nature Structural & Molecular Biology, 22(3), 256–264. https://doi.org/10.1038/nsmb.2959

52 Gaffo, E., Bonizzato, A., Kronnie, G., & Bortoluzzi, S. (2017). CirComPara: A Multi-Method Comparative bioinformatics pipeline to detect and study Circrnas FROM rna-seq data. Non-Coding RNA, 3(1), 8. https://doi.org/10.3390/ncrna3010008

53 Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S. D., Gregersen, L. H., Munschauer, M., Loewer, A., Ziebold, U., Landthaler, M., Kocks, C., le Noble, F., & Rajewsky, N. (2013). Circular rnas are a large class of animal rnas with regulatory potency. Nature, 495(7441), 333–338. https://doi.org/10.1038/nature11928

54 Narci, K., Kahraman, D. C., Koyas, A., Ersahin, T., Tuncbag, N., & Atalay, R. C. (2021). Context dependent isoform specific PI3K INHIBITION Confers drug resistance in hepatocellular carcinoma cells. https://doi.org/10.1101/2021.03.17.435839

55 Expasy is operated by the SIB Swiss Institute of Bioinformatics |Terms of Use. (n.d.). Cellosaurus. Cellosaurus cell line Huh-7 (CVCL_0336). https://web.expasy.org/cellosaurus/CVCL_0336.

56 Expasy is operated by the SIB Swiss Institute of Bioinformatics|Terms of Use. (n.d.). Cellosaurus. Cellosaurus cell line Mahlavu (CVCL_0405). https://web.expasy.org/cellosaurus/CVCL_0405.

57 7 cell Line origins and characteristics. Huh. (2019, November 18). http://huh7.com/.

58 Fatetorbay, R., Son, Y. A., & Tuncbag, N. (2020). Network-bassed Discovery of Molecular Targeted Agent Treatments in Hepatocellular Carcinoma (thesis).

59 Yuzugullu, H., Benhaj, K., Ozturk, N., Senturk, S., Celik, E., Toylu, A., … Ozturk, M. (2009). Canonical Wnt signaling is antagonized by noncanonical Wnt5a in hepatocellular carcinoma cells. Molecular Cancer, 8, 90. https://doi.org/10.1186/1476-4598-8-90

60 Tang, W., Chen, Z., Zhang, W., Cheng, Y., Zhang, B., Wu, F., Wang, Q., Wang, S., Rong, D., Reiter, F. P., De Toni, E. N., & Wang, X. (2020). The mechanisms of sorafenib resistance in hepatocellular carcinoma: Theoretical basis and therapeutic aspects. Signal Transduction and Targeted Therapy, 5(1). https://doi.org/10.1038/s41392-020-0187-x

61 Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2012). STAR: Ultrafast universal RNA-seq aligner. Bioinformatics, 29(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

62 Cheng, J., Metge, F., & Dieterich, C. (2015). Specific identification and quantification of circular rnas from sequencing data. Bioinformatics, 32(7), 1094–1096. https://doi.org/10.1093/bioinformatics/btv656

63 Fu, X., & Renyi, L. (2014). CircRNAFinder: a Tool for Identifying Circular RNAs Using RNA-Seq Data. Proceedings of the 6th International Conference on Bioinformatics and Computational Biology, BICOB.

64 Kahraman, D. C., Kahraman, T., & Cetin Atalay, R. (2019). Targeting PI3K/Akt/mTOR PATHWAY IDENTIFIES DIFFERENTIAL expression and functional role of IL-8 in liver cancer stem cell enrichment. Molecular Cancer Therapeutics. https://doi.org/10.1158/1535-7163.mct-19-0004

65 Yang, J., Nie, J., Ma, X., Wei, Y., Peng, Y., & Wei, X. (2019). Targeting PI3K in Cancer: Mechanisms and advances in clinical trials. Molecular Cancer, 18(1). https://doi.org/10.1186/s12943-019-0954-x

66 Jeck, W. R., Sorrentino, J. A., Wang, K., Slevin, M. K., Burd, C. E., Liu, J., Marzluff, W. F., & Sharpless, N. E. (2012). Circular RNAs are abundant, conserved, and associated with ALU repeats. RNA, 19(2), 141–157. https://doi.org/10.1261/rna.035667.112

67 Zhong, L., Wang, Y., Cheng, Y., Wang, W., Lu, B., Zhu, L., & Ma, Y. (2018). Circular rna circc3p1 suppresses hepatocellular carcinoma growth and metastasis through mir-4641/pck1 pathway. Biochemical and Biophysical Research Communications, 499(4), 1044–1049. https://doi.org/10.1016/j.bbrc.2018.03.221

68 Yu, J., Xu, Q.-guo, Wang, Z.-guang, Yang, Y., Zhang, L., Ma, J.-zhao, Sun, S.-han, Yang, F., & Zhou, W.-ping. (2018). Circular rna csmarca5 inhibits growth and metastasis in hepatocellular carcinoma. Journal of Hepatology, 68(6), 1214–1227. https://doi.org/10.1016/j.jhep.2018.01.012

69 Han, D., Li, J., Wang, H., Su, X., Hou, J., Gu, Y., Qian, C., Lin, Y., Liu, X., Huang, M., Li, N., Zhou, W., Yu, Y., & Cao, X. (2017). Circular rna circmto1 acts as the sponge of microrna-9 to suppress hepatocellular carcinoma progression. Hepatology, 66(4), 1151–1164. https://doi.org/10.1002/hep.29270

70 Yao, Z., Luo, J., Hu, K., Lin, J., Huang, H., Wang, Q., Zhang, P., Xiong, Z., He, C., Huang, Z., Liu, B., & Yang, Y. (2017). ZKSCAN1gene and its related circular RNA (circzkscan1) both inhibit hepatocellular Carcinoma cell GROWTH, migration, and invasion but through different signaling pathways. Molecular Oncology, 11(4), 422–437. https://doi.org/10.1002/1878-0261.12045

71 Fu, L., Yao, T., Chen, Q., Mo, X., Hu, Y., & Guo, J. (2017). Screening differential circular rna expression profiles reveals hsa_circ_0004018 is associated with hepatocellular carcinoma. Oncotarget, 8(35), 58405–58416. https://doi.org/10.18632/oncotarget.16881

72 Huang, X.-Y., Huang, Z.-L., Xu, Y.-H., Zheng, Q., Chen, Z., Song, W., Zhou, J., Tang, Z.-Y., & Huang, X.-Y. (2017). Comprehensive circular rna profiling reveals the regulatory role of the circrna-100338/mir-141-3p pathway in hepatitis b-related hepatocellular carcinoma. Scientific Reports, 7(1). https://doi.org/10.1038/s41598-017-05432-8

73 Wang, B.-gang, Li, J.-shuai, Liu, Y.-feng, & Xu, Q. (2017). MicroRNA-200b suppresses the invasion and migration of hepatocellular carcinoma by DOWNREGULATING RHOA and circrna_000839. Tumor Biology, 39(7), 101042831771957. https://doi.org/10.1177/1010428317719577

74 Xu, L., Zhang, M., Zheng, X., Yi, P., Lan, C., & Xu, M. (2016). The circular rna cirs-7 (cdr1as) acts as a risk factor of hepatic microvascular invasion in hepatocellular carcinoma. Journal of Cancer Research and Clinical Oncology, 143(1), 17–27. https://doi.org/10.1007/s00432-016-2256-7

75 Qin, M., Liu, G., Huo, X., Tao, X., Sun, X., Ge, Z., Yang, J., Fan, J., Liu, L., & Qin, W. (2016). Hsa_circ_0001649: A circular rna and potential novel biomarker for hepatocellular carcinoma. Cancer Biomarkers, 16(1), 161–169. https://doi.org/10.3233/cbm-150552

76 Shang, X., Li, G., Liu, H., Li, T., Liu, J., Zhao, Q., & Wang, C. (2016). Comprehensive circular RNA PROFILING reveals That hsa_circ_0005075, a new Circular Rna Biomarker, is involved in Hepatocellular CRCINOMA DEVELOPMENT. Medicine, 95(22). https://doi.org/10.1097/md.0000000000003811

77 Yu, L., Gong, X., Sun, L., Zhou, Q., Lu, B., & Zhu, L. (2016). The circular RNA Cdr1as act as an ONCOGENE in hepatocellular carcinoma through TARGETING miR-7 Expression. PLOS ONE, 11(7). https://doi.org/10.1371/journal.pone.0158347

78 Xiao, J. (2018). Biogenesis of Circular RNAs. In Circular rnas biogenesis and functions (pp. 41–63). essay, Springer Singapore.

79 Hansen, T. B., Venø, M. T., Damgaard, C. K., & Kjems, J. (2015). Comparison of circular rna prediction tools. Nucleic Acids Research, 44(6). https://doi.org/10.1093/nar/gkv1458

80 Zeng, X., Lin, W., Guo, M., & Zou, Q. (2017). A comprehensive overview and evaluation of circular rna detection tools. PLOS Computational Biology, 13(6). https://doi.org/10.1371/journal.pcbi.1005420

APPENDICES

A. Quality Control Results

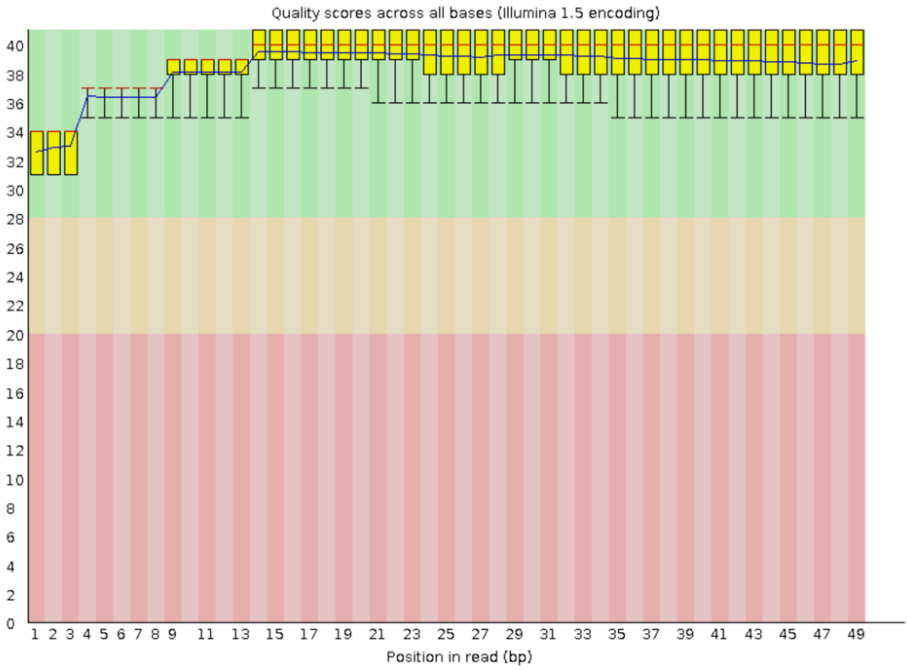

Figure A 1 Per Base Sequence Quality of HUH7 cell line treated with DMSO

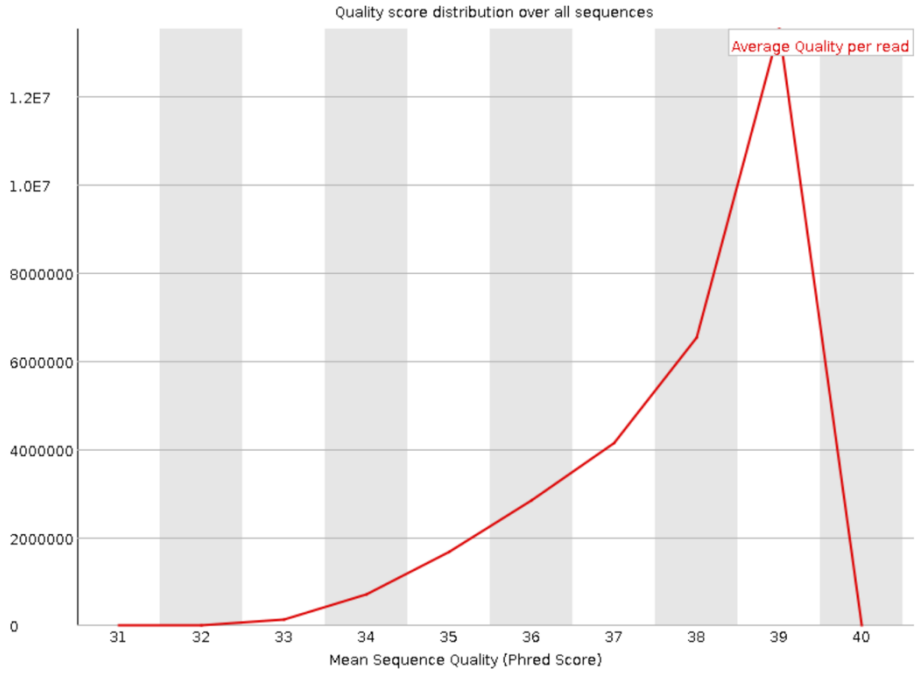

Figure A 2 Per Sequence Quality Scores of HUH7 cell line treated with DMSO

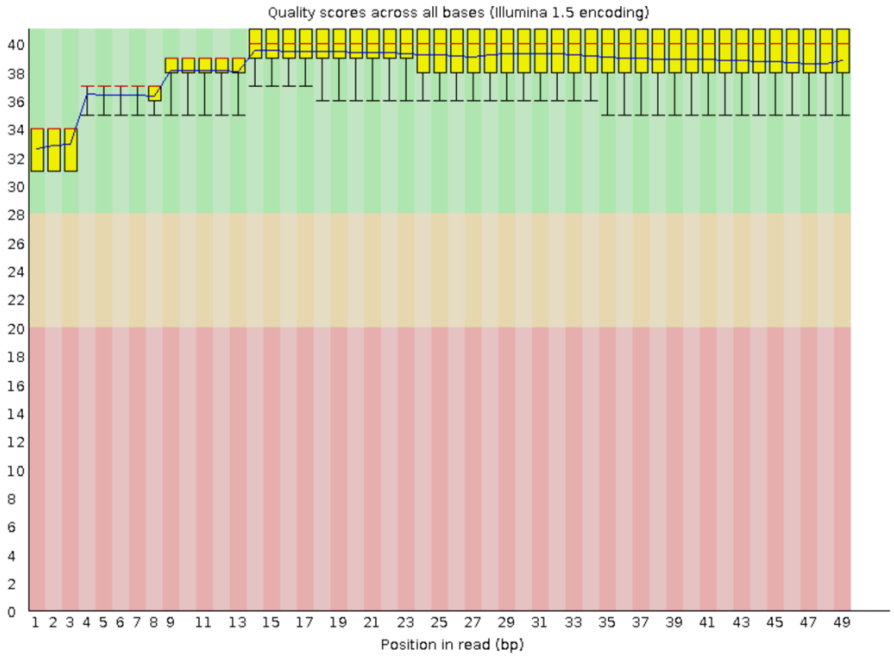

Figure A 3 Per Base Sequence Quality of HUH7 cell line treated with PI3K-β

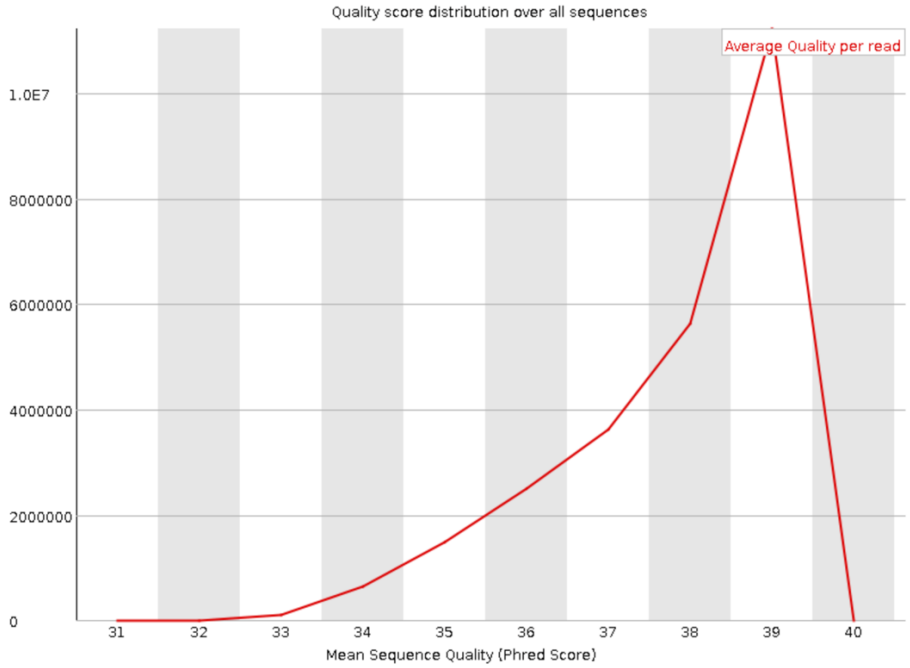

Figure A 4 Per Sequence Quality Scores of HUH7 cell line treated with PI3K-β

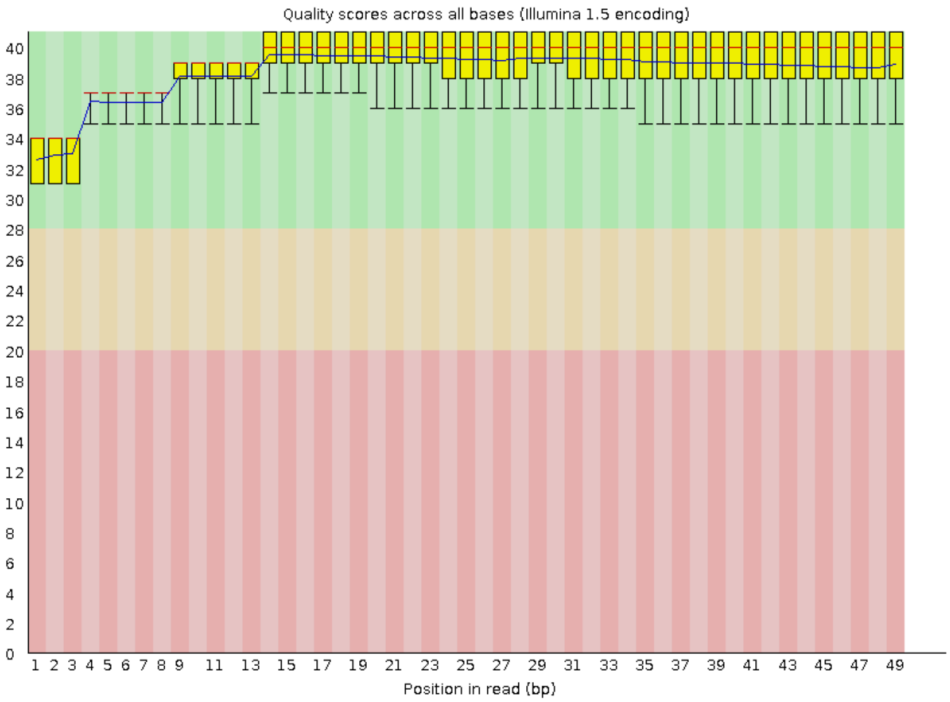

Figure A 5 Per Base Sequence Quality of HUH7 cell line treated with Sorafenib

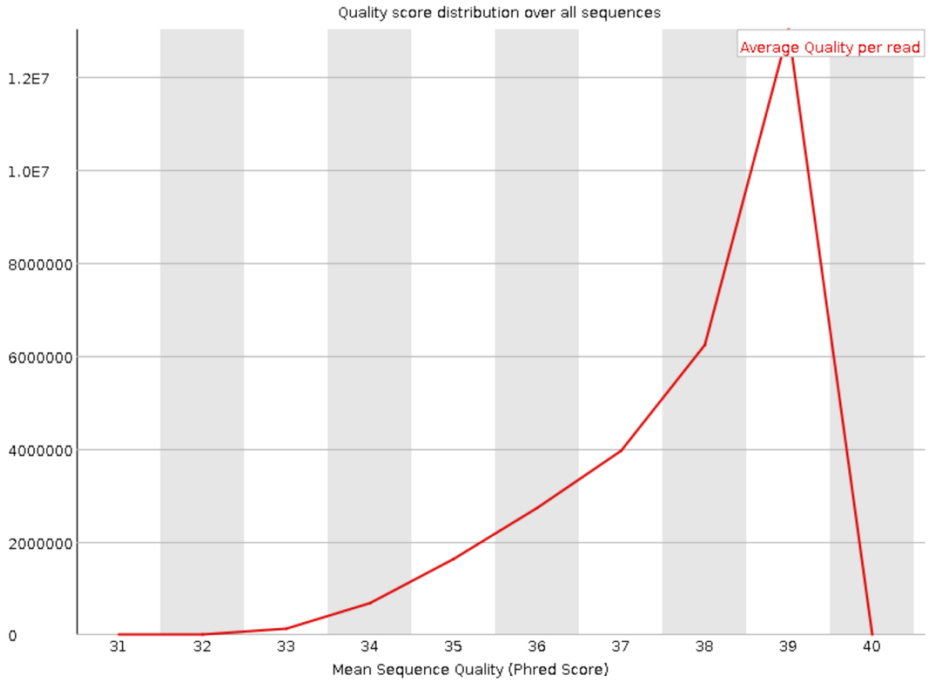

Figure A 6 Per Sequence Quality Scores of HUH7 cell line treated with Sorafenib

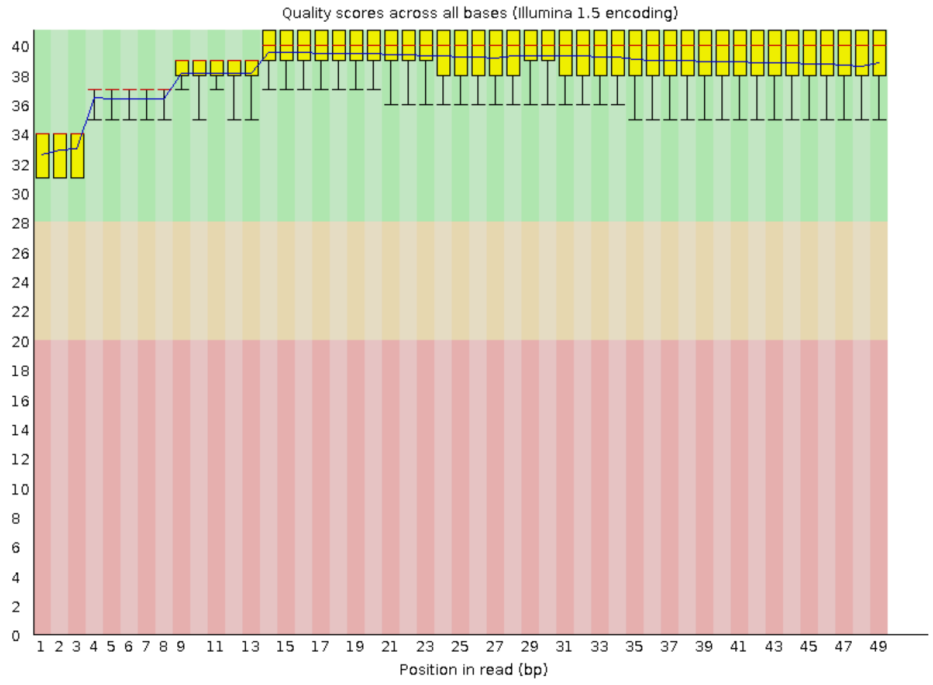

Figure A 7 Per Base Sequence Quality of HUH7 cell line treated with Sorafenib + PI3K-α

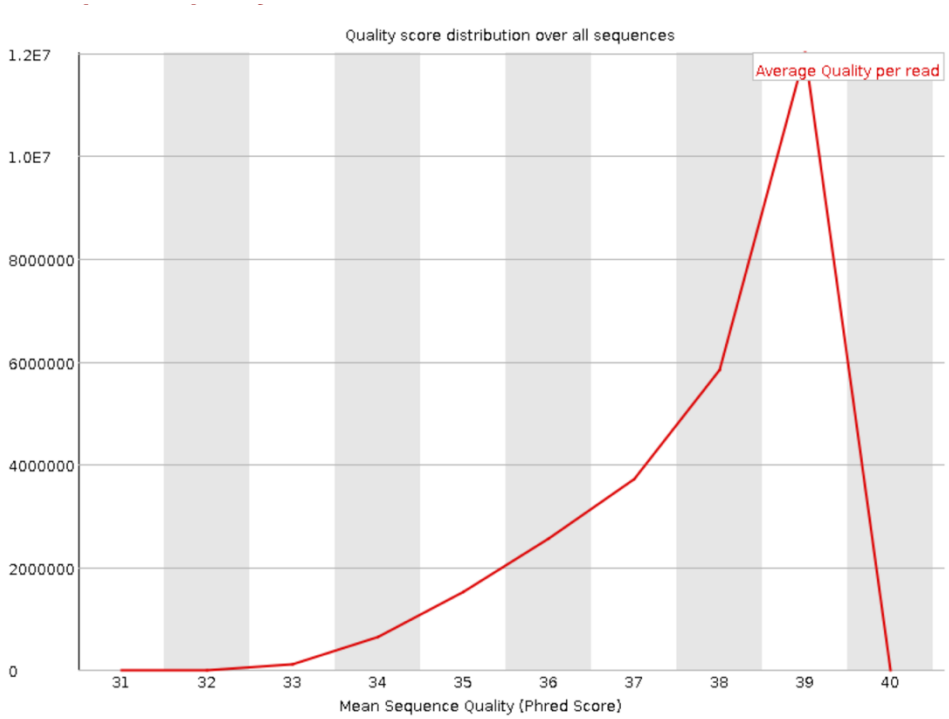

Figure A 8 Per Sequence Quality Scores of HUH7 cell line treated with Sorafenib + PI3K-α

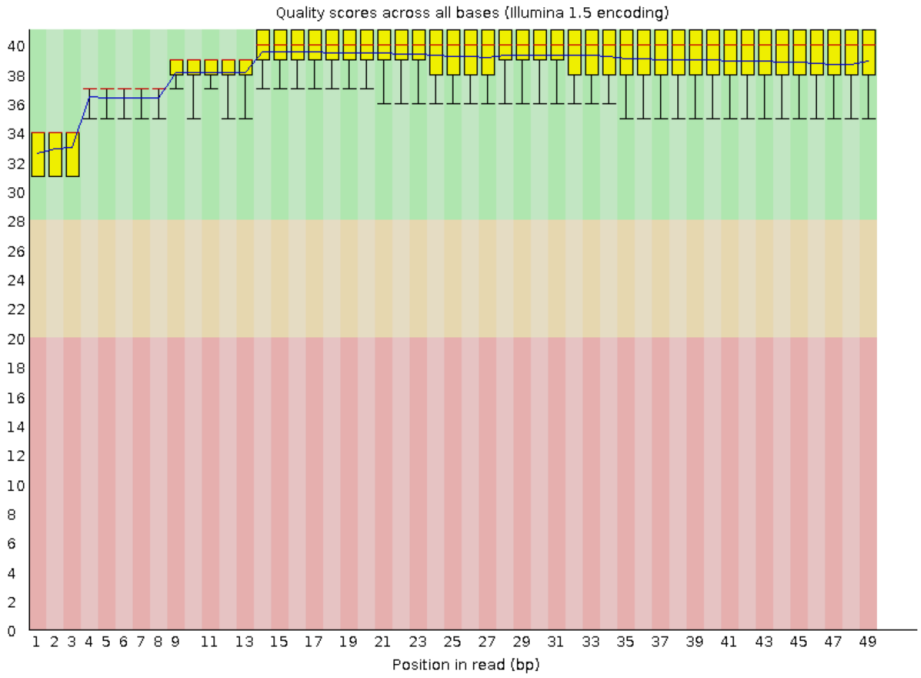

Figure A 9 Per Base Sequence Quality of HUH7 cell line treated with Sorafenib + PI3K-β

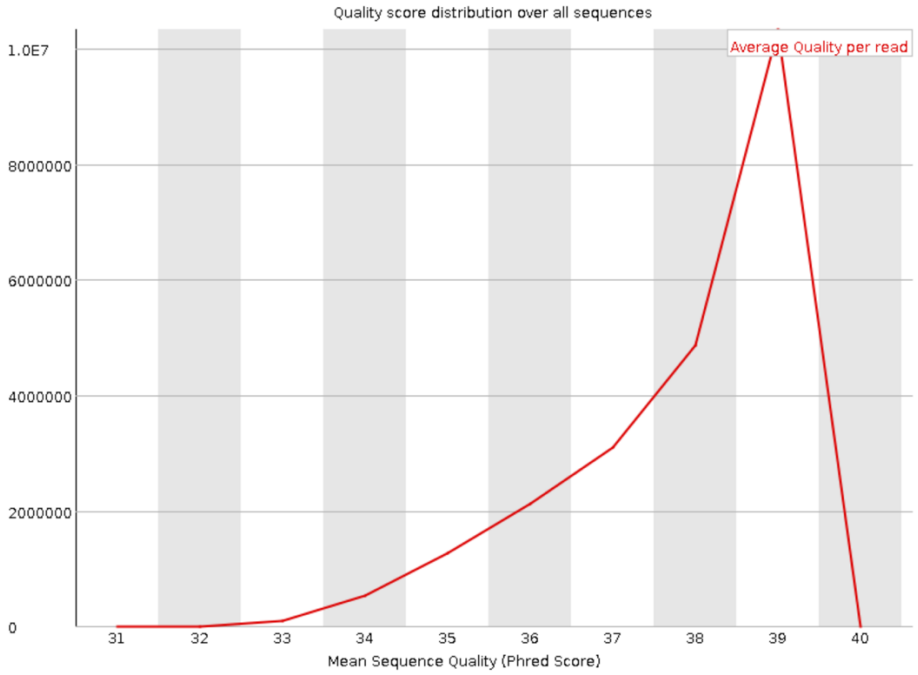

Figure A 10 Per Sequence Quality Scores of HUH7 cell line treated with Sorafenib + PI3K-β

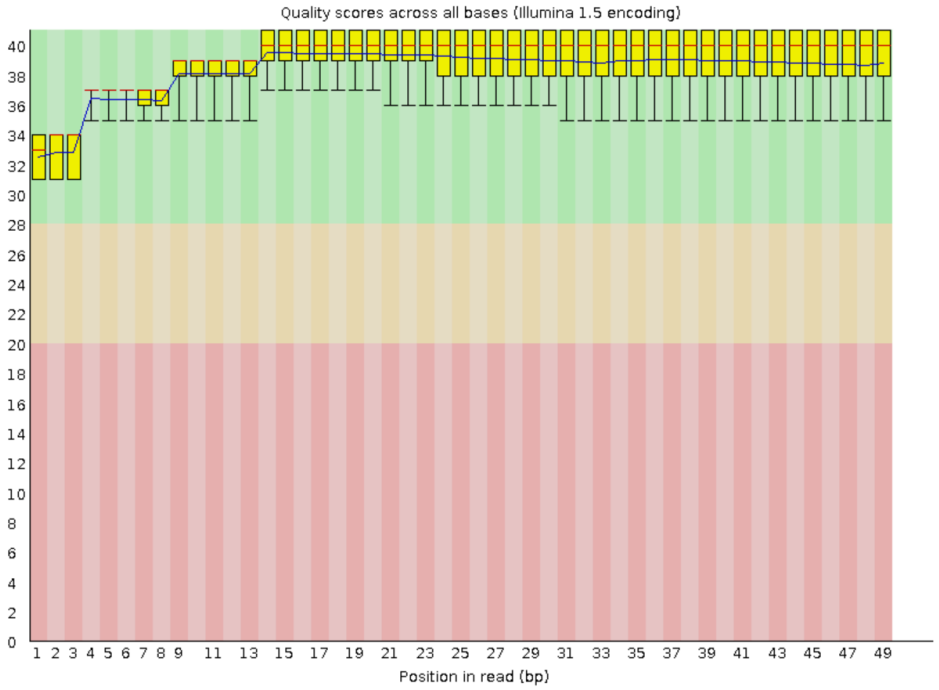

Figure A 11 Per Base Sequence Quality of Mahlavu cell line treated with DMSO

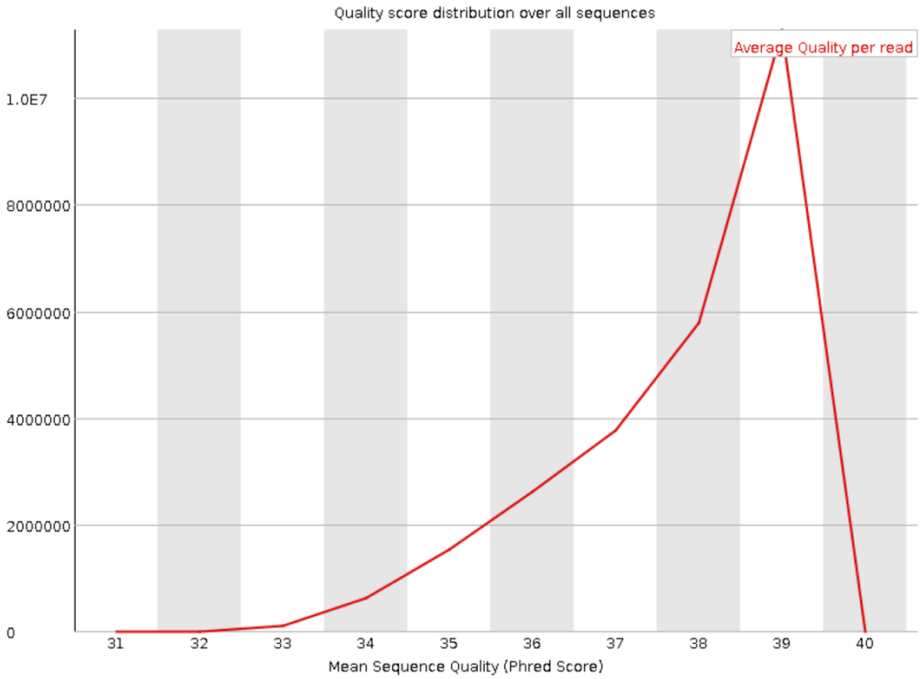

Figure A 12 Per Sequence Quality Scores of Mahlavu cell line treated with DMSO

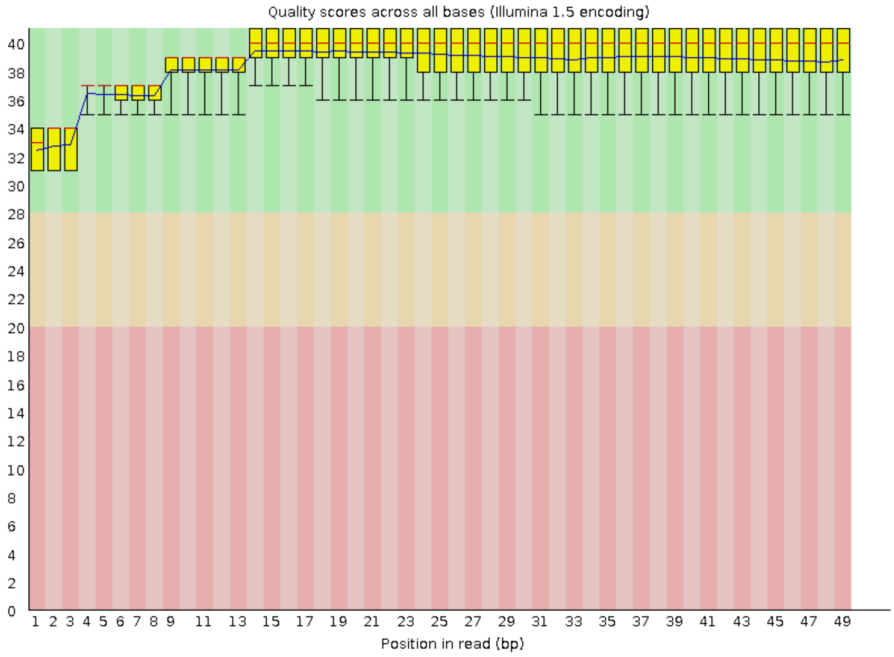

Figure A 13 Per Base Sequence Quality of Mahlavu cell line treated with PI3K-α

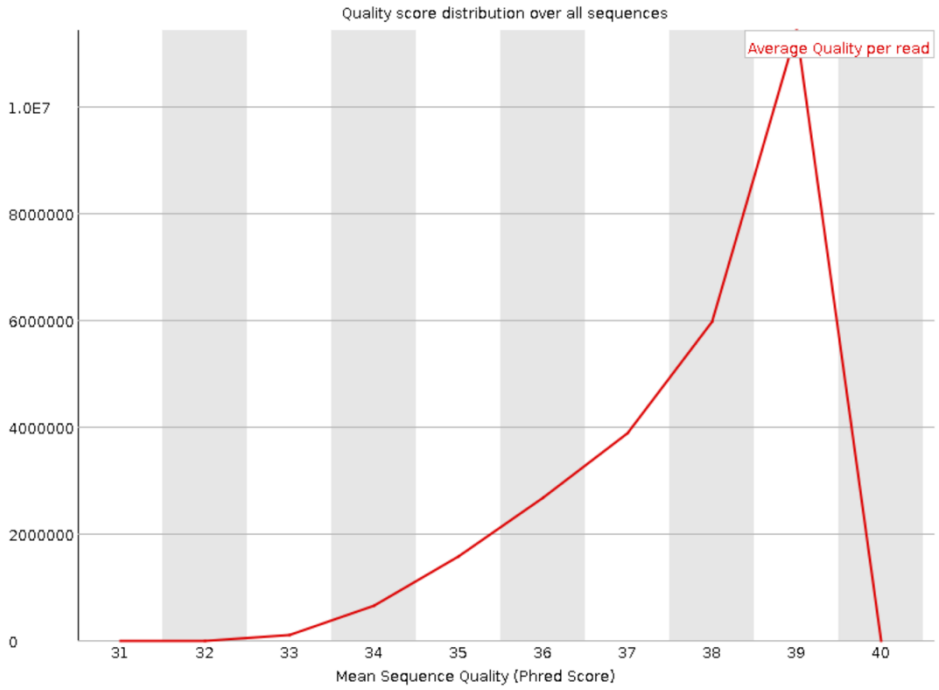

Figure A 14 Per Sequence Quality Scores of Mahlavu cell line treated with PI3K-α

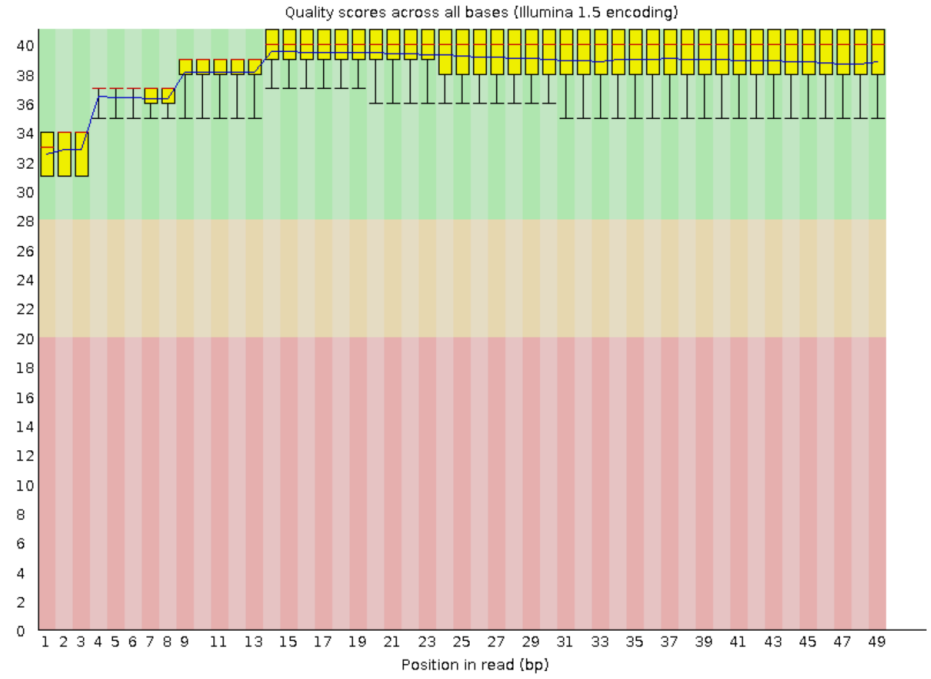

Figure A 15 Per Base Sequence Quality of Mahlavu cell line treated with PI3K-β

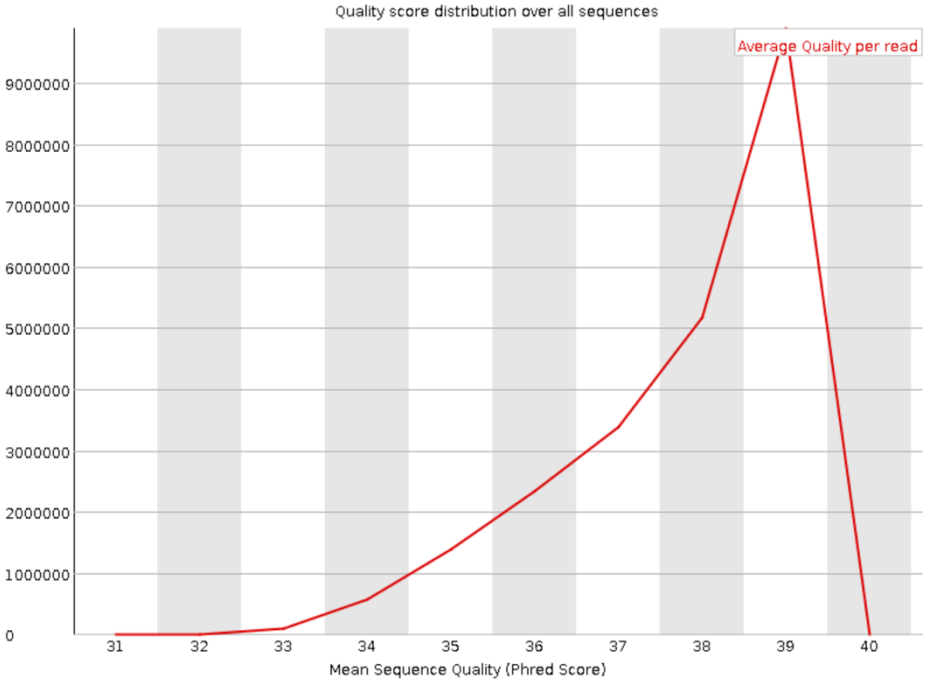

Figure A 16 Per Sequence Quality Scores of Mahlavu cell line treated with PI3K-β

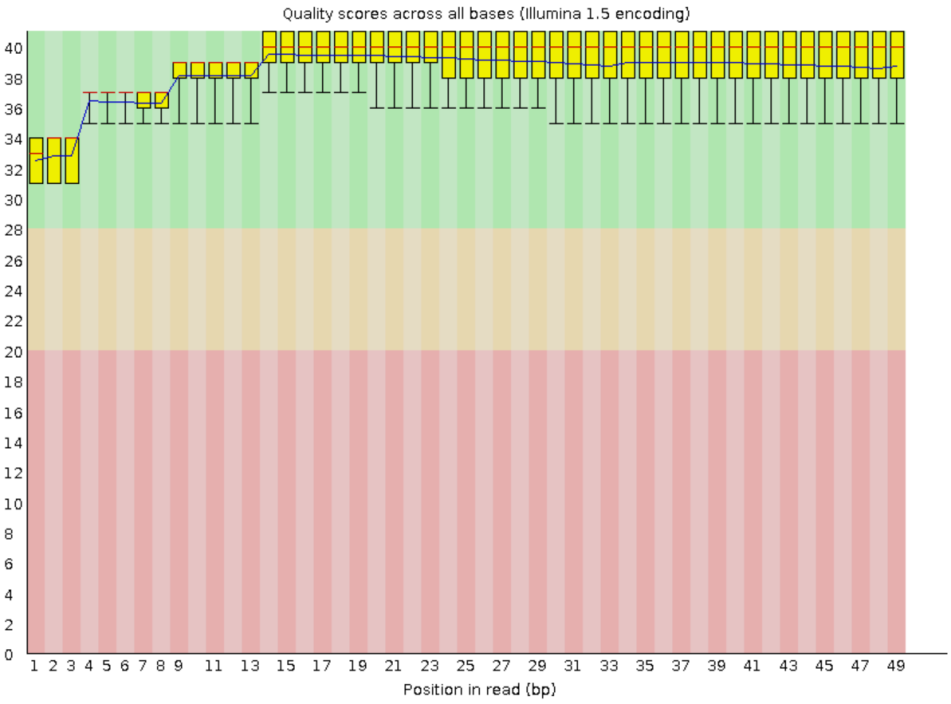

Figure A 17 Per Base Sequence Quality of Mahlavu cell line treated with Sorafenib

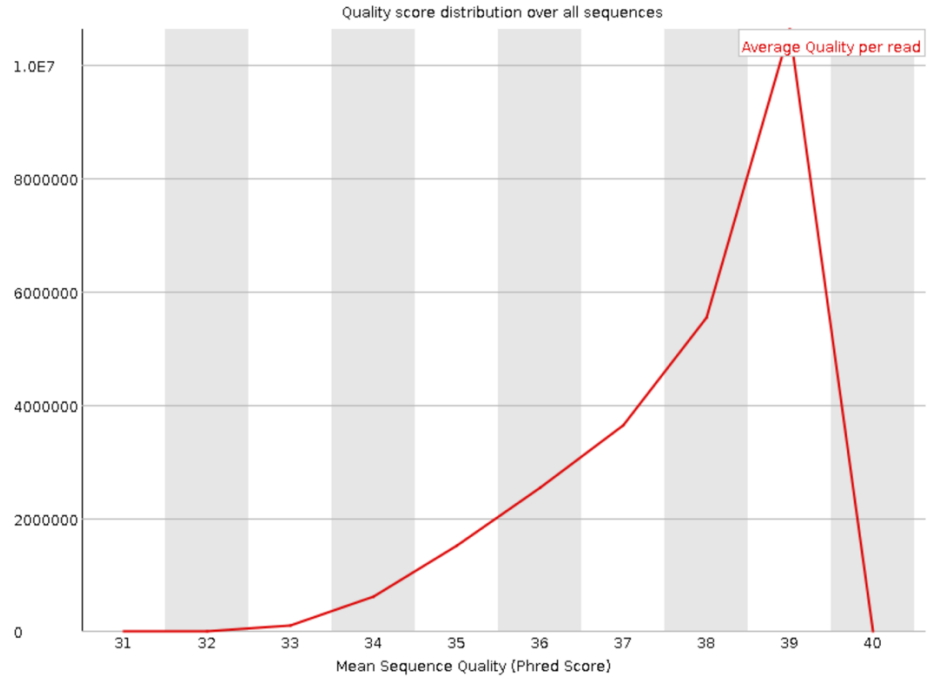

Figure A 18 Per Sequence Quality Scores of Mahlavu cell line treated with Sorafenib

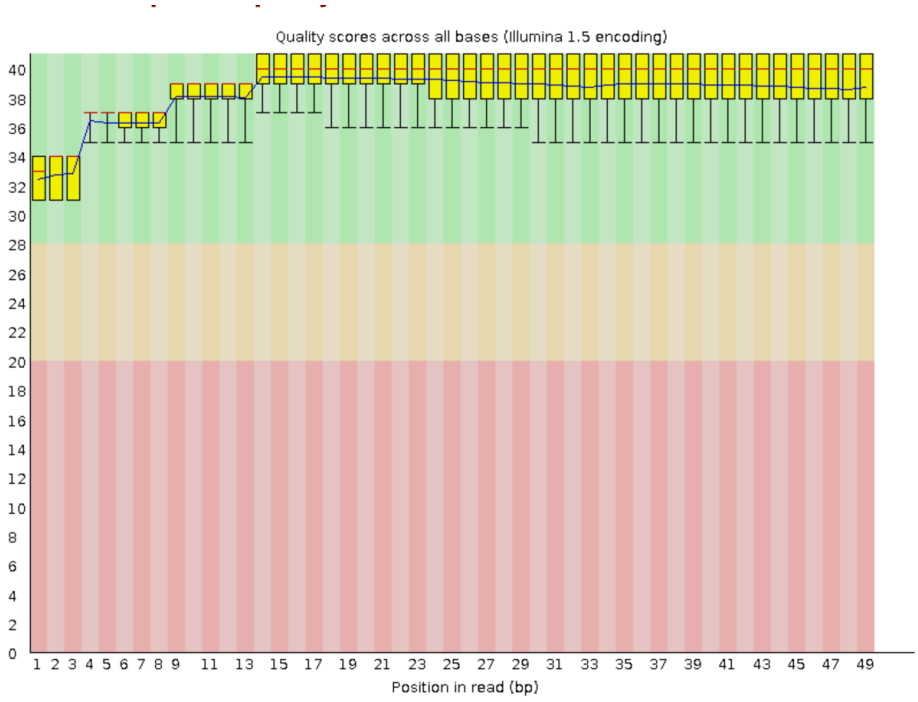

Figure A 19 Per Base Sequence Quality of Mahlavu cell line treated with Sorafenib + PI3K-α

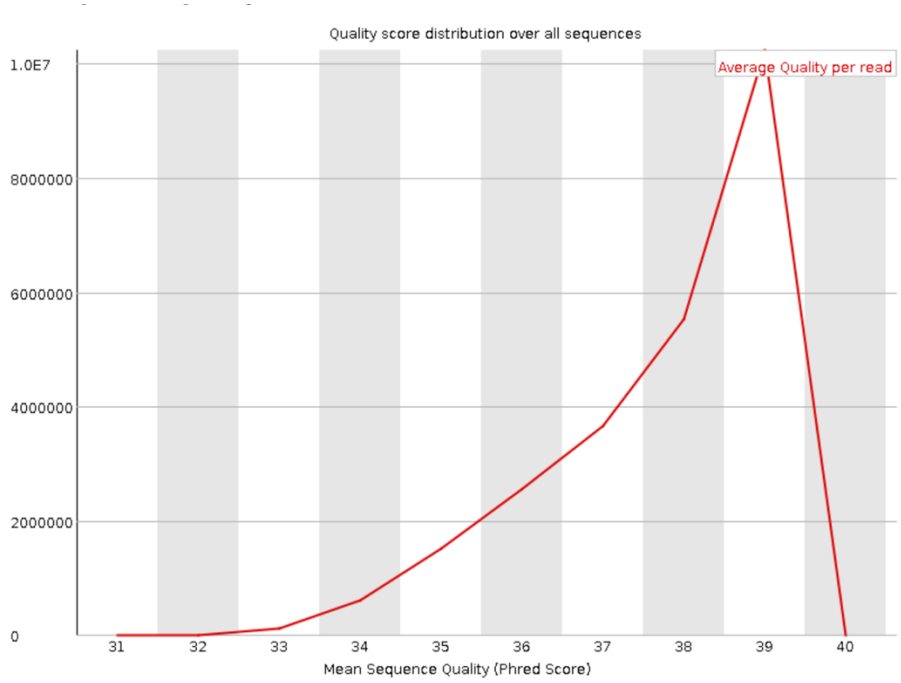

Figure A 20 Per Sequence Quality Scores of Mahlavu cell line treated with Sorafenib + PI3K-α

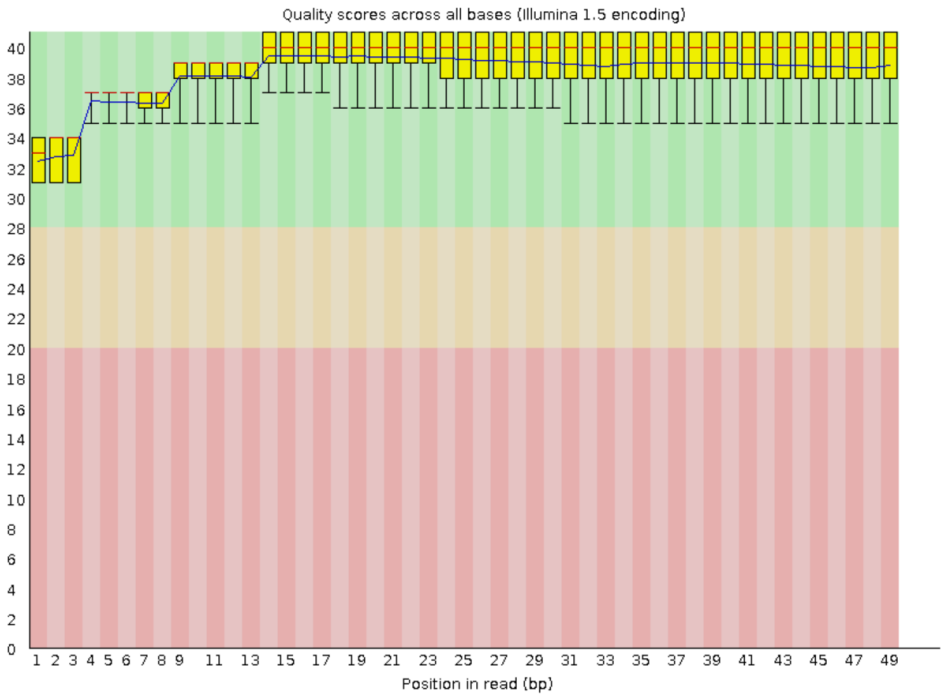

Figure A 21 Per Base Sequence Quality of Mahlavu cell line treated with Sorafenib + PI3K-β

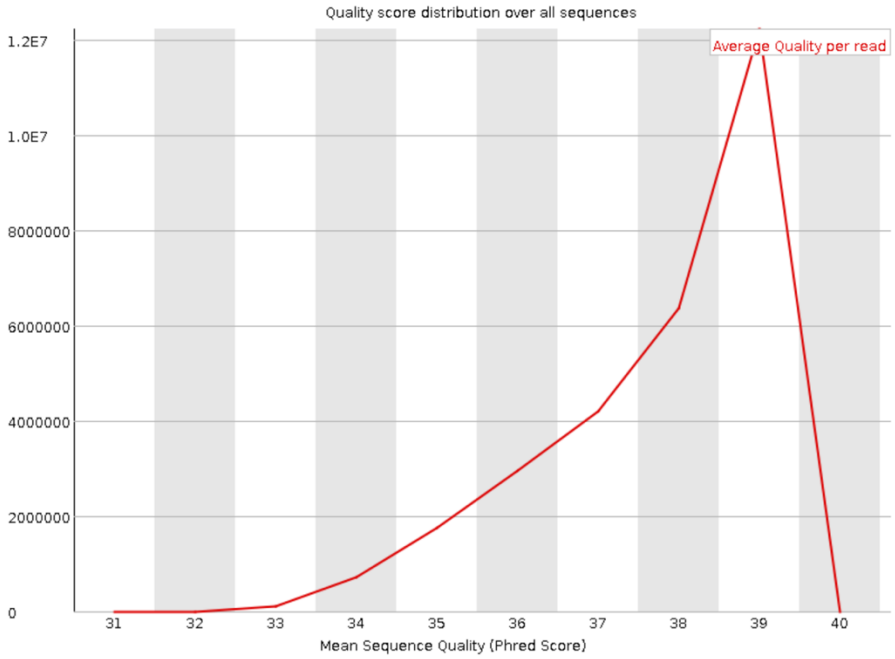

Figure A 22 Per Sequence Quality Scores of Mahlavu cell line treated with Sorafenib + PI3K-β

B. Execution Files

HUH-7 cell line files:

Meta.csv file

file,sample,condition

/home/burcut/RCA_RNASeq2015/raw/clean_reads/H_cells/H-ALPHA/FCC6N26ACXX-HUMoiuRAACRAAPEI-111_L1_1.fq,H_cell_ALPHA,DRUG

/home/burcut/RCA_RNASeq2015/raw/clean_reads/H_cells/H-BETA/FCC6N26ACXX-HUMoiuRAAERAAPEI-113_L1_1.fq.gz,H_cell_BETA,DRUG

/home/burcut/RCA_RNASeq2015/raw/clean_reads/H_cells/H-DMSO/FCC6N26ACXX-HUMoiuRAAARAAPEI-109_L1_1.fq.gz,H_cell_DMSO,NO_DRUG

/home/burcut/RCA_RNASeq2015/raw/clean_reads/H_cells/H-S-ALPHA/FCC6N26ACXX-HUMoiuRAADRAAPEI-112_L1_1.fq.gz,H_cell_S_ALPHA,DRUG

/home/burcut/RCA_RNASeq2015/raw/clean_reads/H_cells/H-S-BETA/FCC6N26ACXX-HUMoiuRAAFRAAPEI-114_L1_1.fq.gz,H_cell_S_BETA,DRUG

/home/burcut/RCA_RNASeq2015/raw/clean_reads/H_cells/H-SOR/FCC6N26ACXX-HUMoiuRAABRABPEI-110_L1_1.fq.gz,H_cell_SOR,DRUG

Vars.py file

```
META            = 'meta.csv'

GENOME_FASTA    =
'/home/burcut/ref_indexes/Homo_sapiens/Ensembl/GRCh37/Sequence/WholeGeno
meFasta/genome.fa'

ANNOTATION      =
'/home/burcut/ref_indexes/Homo_sapiens/Ensembl/GRCh37/Annotation/Archives/
archive-current/Genes/genes.gtf'

CPUS            = '11'

PREPROCESSOR    = 'trimmomatic'

PREPROCESSOR_PARAMS = 'MAXINFO:40:0.5 LEADING:20 TRAILING:20
SLIDINGWINDOW:4:30 MINLEN:35 AVGQUAL:30'

CIRCRNA_METHODS = 'ciri,findcirc,'\

                'circexplorer2_star,dcc,'\

                'circexplorer2_tophat,circrna_finder'


TOGGLE_TRANSCRIPTOME_RECONSTRUCTION = 'False'

FIX_READ_HEADER = 'True'

#MIN_READS = 2 #default


## aligners' custom parameters

HISAT2_EXTRA_PARAMS = '--rna-strandness RF' # stranded libraries

## parameters from CIRI
```

```
BWA_PARAMS = ['-T', '19', '-c', '1']

## parameters from CIRCexplorer2

SEGEMEHL_PARAMS = ['-M','1'] #'-D', '0', '-Z', '20'

TOPHAT_PARAMS = ['--max-multihits', '1']#'--zpacker','pigz'

## parameters used in DCC manual example

STAR_PARAMS = ['--outFilterMultimapNmax', '1',
        '--outSJfilterOverhangMin', '15', '15', '15', '15',
        '--alignSJoverhangMin', '15',
        '--alignSJDBoverhangMin', '15',
        '--seedSearchStartLmax', '30',
        '--outFilterScoreMin', '1',
        '--outFilterMatchNmin', '1',
        '--outFilterMismatchNmax', '2',
        '--chimSegmentMin', '15',
        '--chimScoreMin', '15',
        '--chimScoreSeparation', '10',
        '--chimJunctionOverhangMin', '15']


## pre-computed index and annotation files

GENOME_INDEX =
"/home/burcut/circrna_tools/integrated_tools/CirComPara/test_circompara/whole_
SRA_20_RNA_seq_SE_previous_created_indexes/annotation_indexes/hisat2/geno
me"
```

#SEGEMEHL_INDEX =
"/home/burcut/circrna_tools/integrated_tools/CirComPara/test_circompara/analysis_se/annotation_indexes/segemehl/CFLAR_HIPK3.idx"

BWA_INDEX =
"/home/burcut/circrna_tools/integrated_tools/CirComPara/test_circompara/whole_SRA_20_RNA_seq_SE_previous_created_indexes/annotation_indexes/bwa/genome"

BOWTIE2_INDEX =
"/home/burcut/circrna_tools/integrated_tools/CirComPara/test_circompara/whole_SRA_20_RNA_seq_SE_previous_created_indexes/annotation_indexes/bowtie2/genome"

BOWTIE_INDEX =
"/home/burcut/circrna_tools/integrated_tools/CirComPara/test_circompara/whole_SRA_20_RNA_seq_SE_previous_created_indexes/annotation_indexes/bowtie/genome"

STAR_INDEX =
"/home/burcut/circrna_tools/integrated_tools/CirComPara/test_circompara/whole_SRA_20_RNA_seq_SE_previous_created_indexes/annotation_indexes/star/genome"

GENEPRED =
"/home/burcut/circrna_tools/integrated_tools/CirComPara/test_circompara/whole_SRA_20_RNA_seq_SE_previous_created_indexes/annotation_indexes/genes.genePred.wgn"

LIN_COUNTER = 'ccp' #'dcc'

DCC_EXTRA_PARAMS = ['-fg', '-M', '-Nr', 1, 1, '-F', '-ss']

TESTREALIGN_PARAMS = ['-q', 'median_1']

CE2_PARAMS = ['--no-fix'] #suggested not to set '--no-fix' in real datasets

```
FINDCIRC_EXTRA_PARAMS = ['--best-qual', '40']

FIX_READ_HEADER = 'True'

SAM_SORT_MM = '1G'

#BYPASS=’linear’
```