Institute of Human Genetics
Medical University of Graz
Harrachgasse 21/8 A-8010 Graz
Head: Univ.-Prof. Dr.med.univ. Michael Speicher

Master's Thesis

Fast Detection of Aneuploidy in Cancer with Next Generation Sequencing

to achieve the degree of
Master of Science

Author: Teresa Gerhalter, BSc

Evaluator: Marcel Scheideler, Dipl.-Chem. Dr.rer.nat. Univ.-Doz.

Institute of Molecular Biotechnology
Technical University of Graz
Petersgasse 14 A-8010 Graz

Supervisor: Michael Speicher, Dr.med.univ. Univ.-Prof.
Co-supervisor: Ellen Heitzer, Univ.-Ass.in Mag.a Dr.in rer.nat

1/5/2014

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

…… ……..

date (signature)

ACKNOWLEDGEMENT

First, I want to thank Univ.-Prof. Speicher, the head of the Institute of Human Genetics at Medical University of Graz, who created the possibility of doing my master thesis at his institute. In addition, I want to express my special thanks to Univ.-Ass. Heitzer, who gave me a lot of support during my work in the lab, as well as valuable input and feedback. I also want to extend my thanks to Peter Ulz, who helped me with analysing my data.

Furthermore, I am thankful for my work colleagues, as they introduced me into the lab work at the beginning of the labour.

Last but not least, my supervisor Univ.-Doz. Scheideler, who made it possible for me to write an external master thesis. I am truly thankful for this experience I was able to have.

And of course, I am glad to have friends who gave their time for some proofreading. Thank you!

CONTENTS

ABSTRACT

Genomic instability caused by single nucleotide changes, structural changes of the chromosomes, and gains and losses of the entire chromosomes has been associated with human cancers. This dynamic state of tumour cells causes' aneuploidy and intra-tumour heterogeneity. Circulating plasma DNA can be used as a non-invasive, sensitive and accurate method of diagnosis of aneuploidy. Yet, analysis of those chromosomal copy number changes is associated with higher time and cost requirements [Swarup, Rajeswari (2007); Kinde *et al.* (2012); Leary *et al.* (2012); Forshew *et al.* (2012); Beckman *et al.* (2012)].

Recently, Kinde *et al.* (2012) published an improved sequencing method, named Fast Aneuploidy Screening Test-Sequencing System (FAST-SeqS). This method was originally developed for a prenatal testing of aneuploidy by massively parallel sequencing of maternal plasma DNA. It achieves an increased throughput and decreased cost by defining a number of fragments and amplifying using a single primer pair.

Hence, for analysis of chromosomal copy number changes in tumour cells in this study, this new fast technique was established in order to examine plasma DNA from tumour patients. Plasma DNA samples were obtained from colon (n=1), breast (n=28), and prostate (n=25) cancer patients. Instead of a whole genome library preparation and the complex analysis, the preparation process was simplified using only a defined number of repetitive long interspersed elements (LINE sequences) using a single primer pair followed by massive parallel sequencing. In addition, female and male controls without any indications of cancer were sequenced. Furthermore dilution experiments with four cell lines were sequenced to detect the limit and sensitivity of the new Fast method. The chromosomal copy number status was determined by calculating the z-scores for chromosomes, single chromosome arms and the entire genome.

For all samples with malignant disease chromosomal copy changes were observed. Several aneuploidies in circulating cell free DNA, which were already identified with array comparative genomic hybridization (array CGH) or PlasmaSeq, were confirmed using the genome-wide z-scores.

It was found that chromosomal copy number variations of tumour samples can be detected with the new screening method that analyses only a defined number of LINE sequences. Its advantage over whole genome sequencing is seen in the simple implementation, lower cost, decreased analysis time, and higher throughput. Further studies involving a large number of cancer patients will reveal whether it will perform as well as whole genome sequencing in the clinic.

ZUSAMMENFASSUNG

Störungen in der Stabilität des Genoms werden mit Krebs assoziiert, welche zu Änderungen von einzelnen Nucleotiden, Strukturänderungen der Chromosomen und Verlust oder Zunahme eines ganzen Chromosoms führen. Dieser dynamische Prozess ist die Ursache für Aneuploidie und intratumore Heterogenitäten. Zirkulierende Plasma DNA kann verwendet werden um nichtinvasiv, sensitiv und exakt die Chromosomenveränderungen zu detektieren. Diese Analyse ist jedoch noch mit einem hohen Kosten- und Zeitaufwand verbunden [Swarup, Rajeswari (2007); Kinde *et al.* (2012); Leary *et al.* (2012); Forshew *et al.* (2012); Beckman *et al.* (2012)].

Kinde *et al.* (2012) stellten vor Kurzem eine verbesserte Sequenziermethode namens *Fast Aneuploidy Screening Test-Sequencing System* (FAST-SeqS) vor. Diese Methode wurde ursprünglich für die Prenataldiagnostik entwickelt, um Aneuploidie durch beschleunigtes Sequenzieren durch hochparellelen Einsatz der mütterlichen Plasma- DNA zu detektieren. Mit Hilfe einer begrenzten Anzahl von Fragmenten und deren Vervielfältigung mit einem einzigen Primerpaar wird ein erhöhter Durchsatz bei geringeren Kosten erreicht.

Diese neue Technik wurde bei der Plasma- DNA von Krebspatienten für die Analyse von chromosomalen Kopienzahlvariationen verwendet. Die untersuchten Plasma- DNA- Proben inkludieren Darm- (n=1), Brust- (n=28) und Prostatakrebspatienten (n=25). Anstatt einer genomweiten Bibliothek und deren komplexen Analyse wurde der Sequenziervorbereitungsschritt durch die Verwendung von definierten repetitiven LINE-Sequenzen (long interspersed elements) mittels eines einzigen Primerpaars vereinfacht. Zusätzlich wurden noch weibliche und männliche Kontrollen ohne Krebsdiagnose sequenziert. Um die Sensitivität der neuen, schnelleren Methode zu testen, wurden vier Verdünnungsreihen von Krebszelllinien hergestellt. Der Aneuploidie- Status wurde durch den Z-Score für die einzelnen Chromosomen und Chromosomenarme und zusätzlich für das ganze Genom festgelegt.

Es wurden für alle Proben der KrebspatientInnen und Krebszelllinien Kopienzahlvariationen detektiert. Mehrere Aneuploidien der zirkulierenden zellfreien DNA- Proben, welche schon mittels Array- CGH (Comparative genomic hybridization) und PlasmaSeq festgestellt wurden, konnten durch den genomweiten Z-Score bestätigt werden.

Kopienzahlvariationen von Tumorproben können mit der neuen Screeningtechnik detektiert werden, welche den Vorteil gegenüber der herkömmlichen genomweiten Sequenzierung in der einfacheren Ausführung, geringeren Kosten, kürzerer Analysezeit und höherem Durchsatz hat. Weiterführende Studien mit einer größeren Anzahl von Krebspatienten könnten die weitere Relevanz in der klinischen Diagnostik unterstützen.

1. INTRODUCTION

1.1 HUMAN GENETICS

Every single cell of the 10^{13} cells of an adult human being contains in the nucleus the genetic program with the information for its survival. That information is passed to the daughter cells with every cell division. Approximately 200 diverse types of cells differentiate due to genetic signals and possess diverse cellular and molecular functions. A cell is the smallest living unit on earth that is separated by a membrane. The integrity of the genetic program needs to be preserved in a reliable manner, yet it should provide flexibility in formulating a wide variety of responses to different environmental changes on a long-term basis [Klug, Cummings (2003)].

Deoxyribonucleic acid (DNA) is a long, ladder-like molecule that forms a double helix and it carries the whole genetic information. This molecule is organised with proteins into a complex structure called chromosomes. A healthy human cell holds 23 pairs of chromosomes; from those chromosomes one pair determines the gender of the human. Females contain one pair of X chromosomes while males have one X chromosome and one Y chromosome. The coiled pieces of DNA contain many genes and regulatory elements, which both encode the big majority of an individual's genetic information.

The central dogma of molecular genetics describes the process of how the genetic information is passed within the cell: DNA is copied to RNA (Ribonucleic acid) which directs the synthesis of proteins. DNA is a nucleic acid and contains four different nucleotides. Those nucleotides consist of one of four bases- adenine (A), guanine (G), thymine (T) or cytosine (C). The bases determine the genetic code and have a complementary relationship between A and T and between G and C. The Watson-Crick Model describes the DNA structure in the following way: two antiparallel polynucleotide chains are coiled around a central axis, forming a double helix. During denaturation of the double-stranded DNA (dsDNA) the hydrogen bonds of the bases, which holds the two strands of polynucleotides together, break, and the DNA strands separate. The second category of nucleic acids is RNA, of which the best known members- messenger RNAs (mRNA) - facilitate the information flow from DNA to proteins, which are their end products. However, the vast majority of the genome, the entirety of an individual's genetic information, does not encode protein coding genes. Less than 5 % of the human genome make up approximately 20,000 genes, which are translated into proteins [Klug, Cummings (2003)]. A large amount of DNA sequences in most organisms seems to be non-coding. However, this portion varies greatly among species including for example over 98 % of the human genome but only about two % for bacteria [Elgar, Vavouri (2008)]. Most of the noncoding sequences including repeat sequences,

telomeres, introns, and pseudogenes, have important biological functions such as transcriptional and translational regulation.

1.1.1 Repetitive sequences

Repetitive sequences unite all DNA sequences in the genome that are present in multiple copies. Almost half of the human genome consists of repetitive DNA sequences [Jurka *et al*. (2007)], which are subdivided into two classes (Figure 1). The first is the tandem repeat, also called satellite, which consists of an array of DNA sequences immediately adjacent to each other. Satellites are formed by recombination of replication in situ [Rodić *et al*. (2013)].

The second much bigger class consists of interspersed repeats, which are DNA fragments inserted randomly throughout the genome rather than appear in tandem including long terminal (LTR) repeats, non-LTR retrotransposons and other repeats. The interspersed repeats have an upper size limit of 20 to 30 kilobases (kb) and are generally inactive, because mostly they are incomplete copies of transposable elements [Jurka *et al.* (2007)]. Taken all intersperse repetitive sequences together, they account for almost 50 % of the human genome. Those repeats are classified into DNA repeats and RNA transposons including LTR repeats and non-LTR retrotransposons. Transposons, also known as mobile DNA or "jumping genes", occur across all eukaryotic genomes differing in their abundance and diversity. They are divided into two classes based on the mechanism of their spread. DNA transposons are segments of DNA that move by reproducing and inserting in the genome using a cut-and-paste mechanism.

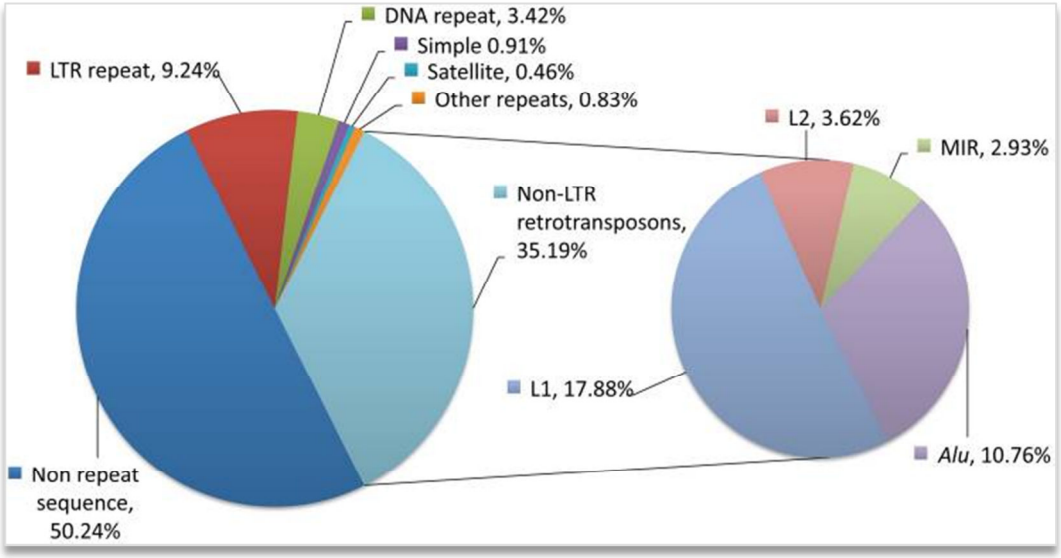


Figure 1: Contribution of different repetitive sequences in human genome [Rodić et al. (2013)].

On the other hand, non-LTR retrotransposons and LTR repeats are transposed through an RNA intermediate by a copy-and-paste mechanism. Both types show autonomous and non-

autonomous variants. The autonomous elements encode all enzymes necessary for their transposition and are therefore self-sufficient. Whereas, non-autonomous elements can only move their genomic DNA copies from one location to another by borrowing the necessary proteins from the autonomous elements [Hancks, Kazazian (2012), Jurka *et al.* (2007)].

Once thought to be random junk DNA, sequencing of whole genomes has generated more information about the origin and genomic impact of the non-coding DNA. The majority of mobile elements represent a relic of previously active elements. At the present only one subclass from the retrotransposons is seen as active in humans [Roy-Engel, Astrid, M. (2012)].

1.1.2 Retrotransposons- LINE and SINE

Retrotransposons are again subdivided into two families based on the presence or absence of long terminal repeats (LTR). Although they both use RNA as an intermediate, their mechanism of insertion differs significantly. Only elements of the non-LTR family are considered to be currently active in humans, namely Long Interspersed Element-1 (LINE-1 or L1) and its non-autonomous partners for example SVA and Alu. More than 30 % of the human genome is derived indirectly or directly from the activity of LINE-1. This autonomous element is capable of transposing a copy of its own RNA or other RNAs. Furthermore, LINE-1 plays a role in pseudogene formation, which makes the gene lose the protein-coding ability, due to retrotransposition of cellular mRNA. Approximately half a million LINE-1 elements are in the entire human genome, but only about 80 to 100 LINE-1 elements are active, because the rest was inactivated in the past through point mutations, truncations, or rearrangements [Hancks, Kazazian (2012)].

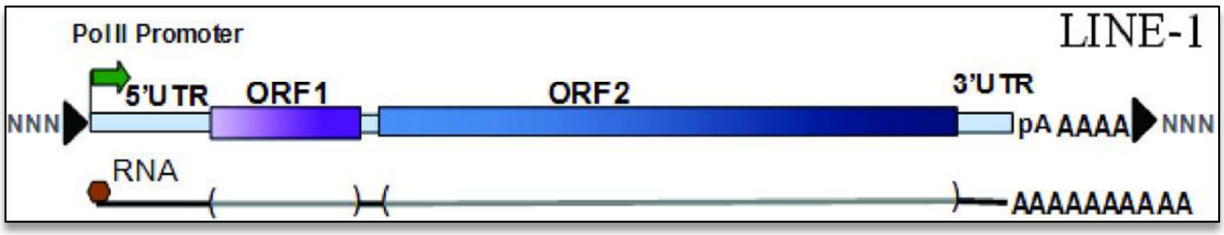


Figure 2: Diagrammed representation of the LINE-1 element with the expected transcript [Roy-Engel, Astrid, M. (2012)]. The transcription start side is indicated by the green arrow.

LINE-1 element belongs to the family of non-LTR retrotransposons. The element is about 6000 bp long with two open reading frames (ORF), ORF1 and ORF2, which are separated by a short spacer (Figure 2). The origin and the exact role of the ORF1 protein are to date unknown. The ORF2 sequence on the other hand encodes a protein with an endonuclease and a reverse transcriptase activity. Endonuclease capability allows cleaving the

phosphodiester bond within a polynucleotide chain and the reverse transcriptase activity enables the translation from an RNA template to the complementary DNA.

An internal RNA polymerase promoter is located in the 5'- untranslated region (UTR) referring to an untranslated coding sequence involved in the transcription regulation. The self-mobilize ability is therefore preserved in the element to be passed on to new inserts. In addition, the LINE-1 features also an antisense promoter located downstream of the RNA polymerase promoter. The ORF2 region is followed by a 3'region containing the own polyA signal and an A-tail.

Another typical non-LTR retrotransposons is the short interspersed element (SINE), which is typical non autonomous. An example for a SINE, which is approximately 300 bp in length, is the Alu element. The Alu and its family members demonstrate no coding capacity and depend on the LINE-1 to obtain the protein machinery needed for retrotransposition. Alu is one of the most active SINE elements with more than one million copies in the human genome. All retrotransposons are not distributed randomly in the genome; however recombination between the chromosomes is believed to play a role in the redistribution [Roy-Engel, Astrid, M. (2012)].

The impact of retrotranspositions remains unclear. Retrotransposons act as a driving force in the evolution of epigenetic regulation and have an impact on genomic stability and evolution. In the human genome approximately 8000 pseudogenes have been identified [Hancks, Kazazian (2012)], which have been processed via retrotransposition of the cellular mRNA (messenger ribonucleic acid). Processing of pseudogenes that influence the appearance of the individual, so called polymorphic pseudogenes, could contribute to the phenotype. Furthermore, 50 to 100 protein-coding genes in the mammalian genome derive from coding sequences of DNA transposons and retrotransposons [Jurka *et al.* (2007)]. Finally, a lot of microRNA genes are thought to evolve from transposon elements. In 1984, McClintock already concluded "that stress, and the genome reaction to it, may underlay many formations of new species", when she noticed, that transposable elements (TE) can be activated quickly due to abiotic stress [McClintock (1984)]. Increased research confirms the thesis that populations benefit from mutations due to activation of TE when they have to deal with extinction. On the other hand, retrotransposition can also cause an insertion that triggers single-gene disease. LINE-1 insertions for instance can drive tumourigenesis by silencing tumour suppressor genes or activating oncogenes. Scientists noticed LINE-1 retrotransposition in colon, lung, prostate, and ovarian carcinomas. For patients suffering from some cancer types like ovarian and lung carcinomas, a high level of LINE-1 RNA can be detected, which is related to poor clinical features [Rodić *et al.* (2013)]. In other cases, the LINE-1 transcription can be induced by drugs and can function as a potential suppressor by inhibiting transcripts of oncogenes. To sum up, the role of LINE-1 in cancer is complex, it is related with tumour suppression as well as with cancer progression [Rodić *et al.* (2013)]. However, most insertions are considered to be benign [Hancks, Kazazian (2012)].

1.2 CANCER GENETICS

Cancer affects fundamental aspects of cellular functions such as DNA repair, cell proliferation, programmed cell death (apoptosis), differentiation, and cell-cell contact. Human tumours emerge during a multistep process, in which genes controlling cell proliferation and metastasis mutate. Affected cells exhibit high levels of genomic instability, which leads to loss of control over how the cells spread and invade the environment [Klug, Cummings (2003)]

The development of cancer is driven by mutations in proto-oncogenes and tumour suppressor genes. However, the number of mutated genes varies widely depending on the type of tumour. Tumours located for example in colon, breast, brain, and pancreas show an average of 33 to 66 mutated genes. The majority (95 %) of these mutations are caused by a single base substitution in the DNA [Vogelstein *et al.* (2013)]. The rest of the mutations are deletions or insertions of more than one base. Nevertheless, "most human cancers are caused by two to eight sequential alterations that develop over the course of 20 to 30 years [Vogelstein *et al.* (2013)]." Environmental agents such as chemicals, ultraviolet light, ionizing radiation, tobacco smoke, and viruses have the potential to damage the DNA and being carcinogenic. But also diet, specially alcohol, and chronic inflammation is implicated in the development of cancer [Klug, Cummings (2003)].

Another aspect is that only approximately one per cent of cancers is associated with germ-line mutations and is therefore inherited. This is another important difference between cancer and other genetic disease in addition to the number of mutations described above.

Cancers differ in their growth rates, invasiveness, and their age of onset. However, all cancer cells show common ground in abnormal cell growth and division (cell proliferation), and in their ability to spread and invade other parts of the body (metastasis) [Klug, Cummings (2003)].

1.2.1 TUMOUR GROWTH AND REGULATION

Genome instability, which is a feature of cancer, produces the genetic diversity accelerating their acquisition, and inflammation. Tumours are complex tissues consisting of a variety of different cell types that show the ability to activate normal cells, which shape a tumour-associated stroma and start to contribute to tumour genesis. Hanahan, Weinber (2011) defined the hallmarks of cancers, including sustaining proliferative signalling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, and activating invasion and metastasis. Other potential hallmarks are reprogramming of energy metabolism and evading immune destruction (Figure 3). Those characteristics play a certain

role in the transformation of normal cells into a malignant state and contribute to the multistep process of tumour growth and metastatic dissemination.

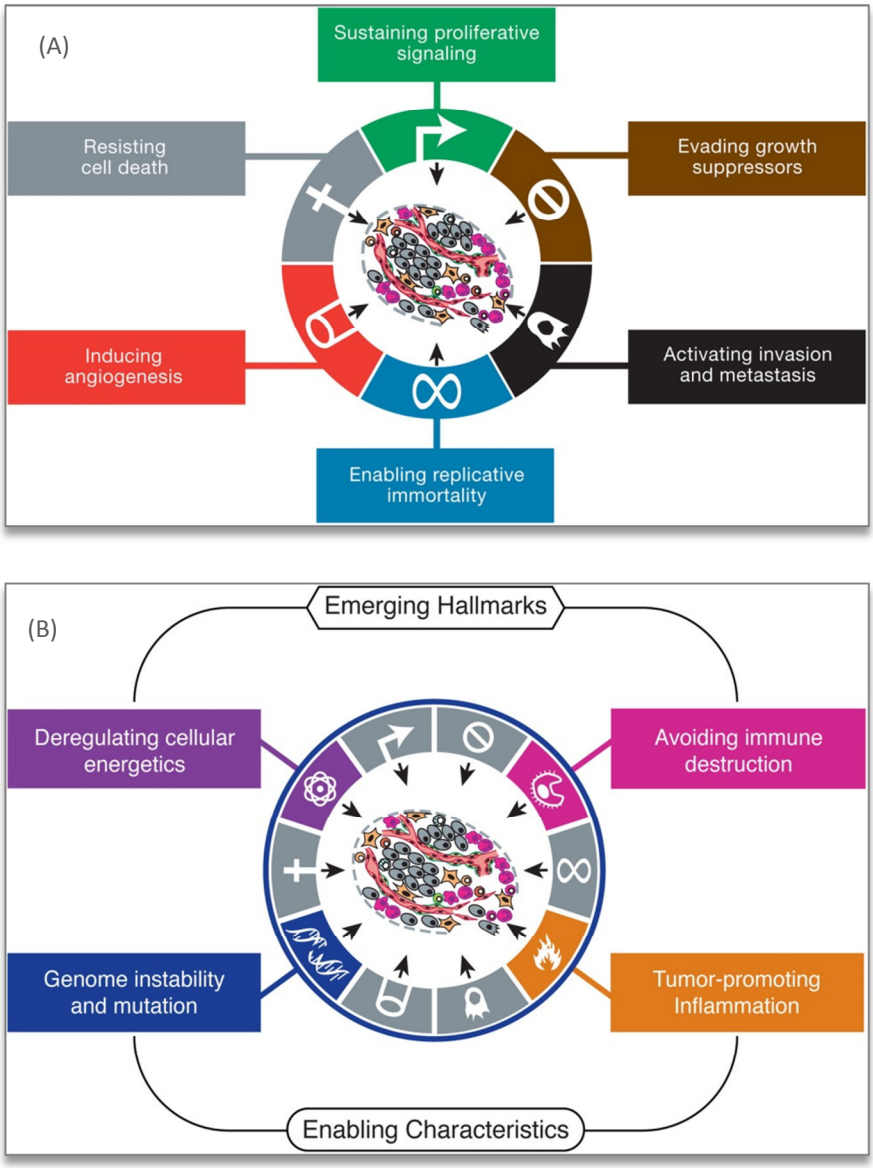


Figure 3: (A) The hallmarks of cancer, (B) Emerging hallmarks of cancer and enabling characteristics [Hanahan, Weinber (2011)].

Normal cells regulate growth, architecture, number, and function with growth-promoting signals. The careful regulation of signal production and distribution ensures a normal proliferation. Cancer cells are able to sustain proliferative signalling in different ways: they emit autocrine and paracrine signals to produce growth factors. Either they regulate their own growth factor ligands, and receptors, or cancer cell stimulate normal cells within the stroma surrounding the cell. In addition, cancer cells need to avoid the negative regulation of cell proliferation.

Usually, programmed cell death, so called apoptosis, is triggered by physiologic stress such as tumourigenesis or anticancer therapy. It acts as a natural barrier, which cancer cells

mostly are able to circumvent. Most common strategies to limit apoptosis are the loss of *TP53* tumour suppressor function, increase of survival signals, and anti-apoptotic regulators [Hanahan, Weinber (2011)]. *TP53* among other tumour suppressor genes limit cell growth and need therefore to be inactivated in order to proliferate at an elevated level. The TP53 protein can interrupt the cell-cycle progression and even trigger apoptosis if the genome is damaged above certain extent. More than half of all cancer patients show a mutation in the *TP53* gene resulting in an inactivation of the TP53 protein [Olivier *et al.* (2010)]. Therefore, inactivation of these tumour suppressor genes among other mutations leads to tumour development.

In addition, cancer cells possess unlimited replicative potential in contrast to normal cells, which undergo only a limited number of growth-and-division cycles. This transition is called immortalization, a process to escape cell death and senescence, which describes an irreversible, non-proliferative state. In normal cells the ends of the chromosomes are protected by telomere sequences shortened with every cell division. However, the enzyme telomerase can maintain the length of the telomere sequences allowing the cells continue to divide. In normal cells telomerase levels are very low and not enough to sustain telomere length of chromosomes. Whereas in cancer cells, especially malignant ones, telomerase is expressed at a significant high level maintaining the telomere length, thus conferring the capability for unlimited proliferation [Blackburn (2005)].

Tissues require nutrients and oxygen for viability and growth as well as the ability to remove metabolic wastes and carbon dioxide. The growing and spreading of vessels (angiogenesis) that addresses the need is a temporary physiologic process in humans triggered usually for wound healing and female reproductive cycling. In cancer cells this angiogenic process is activated permanently in order to satisfy the demand of continuous cell growth and proliferation. Using these new and old vessels, tumour tissues can also spread their cells into the whole body, which is the most frequent cause of death in connection with cancer [Hanahan, Weinber (2011)]. The complex interactions and regulatory mechanism of invasion and metastasis are still not fully understood. It starts with a local invasion of cancer cells, followed by intravasation, which refers to the invasion through the basal membrane into blood or lymphatic vessels. The cancer cells are transported via the vessels, and then settle into distant tissue (extravasation). Only a few cells are capable to form small nodules of cancer cells and to finally grow into macroscopic tumour which is termed as colonization.

Those six defined hallmarks of cancer allow cancer cells to survive, proliferate, and disseminate. Two characteristics facilitate acquisition of those hallmarks; first, genome instability and mutations causing genetic alterations that drive tumour progression. The second enabling characteristic includes the inflammatory state promoted by tumours. Cancer tissues are infiltrated by immune cells and this was thought to reveal the immune system's effort to eliminate the tumours. However, inflammation can support the tumour by supplying growth factors, survival factors, pro-angiogenic factors and many more. In

addition, some molecules released by inflammatory cells are mutagenic for cancer cells, which increase the process of genetic evolution.

Yet, new hallmarks are discovered to play an important role in cancer development. One includes the capability to deregulate cellular energetics in order to support proliferation. Cancer cells can reprogram energy metabolism in order to run cell growth and division. Another emerging hallmark involves active evasion from immune destruction. The immune system constantly monitors cells and eliminates most of the developing cancer cells. As tumours occur, they were able to circumvent detection by cells of the immune system or to reduce the elimination. The hallmarks of cancer and their enabling characteristics give a framework for better understanding the complexity of cancer [Hanahan, Weinber (2011)].

1.2.1 GENOME INSTABILITY AND MUTATION TIMING

All cells from the body arise from the same zygote; therefore they should share the identical genome. But many studies to date reveal an existing genomic variation between different tissues and cells [O'Huallachain *et al.* (2012); Gerlinger *et al.* (2012); Zong *et al.* (2012)]. Many of these genomic changes affect genes with potential hotspots for genomic variation [O'Huallachain *et al*. (2012)].

There are approximately 140 genes revealed to date, which promote tumourgenesis, when they mutate. Furthermore, studies have discovered more driver mutations in tumour suppressor genes than in oncogenes. Other genes show altered epigenetic mechanism including DNA methylation, histone modification, and silencing by microRNA. However, genes with epigenetic changes that cause selective growth advantage are still difficult to identify [Vogelstein *et al.* (2013)].

If a mutation grants a selective growth advantage to the cell, it is termed as driver mutation. On the other hand passenger mutations do not change a cell's condition, but are acquired by driver mutations and appear together with the altered driver genes [Bozic *et al.* (2010)].

Two to eight mutations in these driver genes occur in a typical tumour (Figure 4), although in some tumours such as medullablastomas (brain tumour) the number of driver gene mutations is lower (zero to one). Breast cancer on the other side exhibit more mutations in onco- and tumour suppressor genes. All driver genes are classified into twelve different signalling pathways, which regulate three core cellular processes: cell fate, cell survival, and genome maintenance [Vogelstein *et al.* (2013)]. Although individual tumour types differ with respect to its genetic alterations, they affect those similar cellular processes. First, the balance between differentiation and division of the cell is disturbed resulting in an increased growth advantage. Cancer cells have the capability to proliferate under lower nutrient concentration and are prone to toxic substances and errors during DNA replication avoiding apoptosis. Some genes are involved in several pathways regulating the three core cellular

processes, and therefore mutated genes can have an impact on numerous pathways [Hanahan, Weinber (2011)].

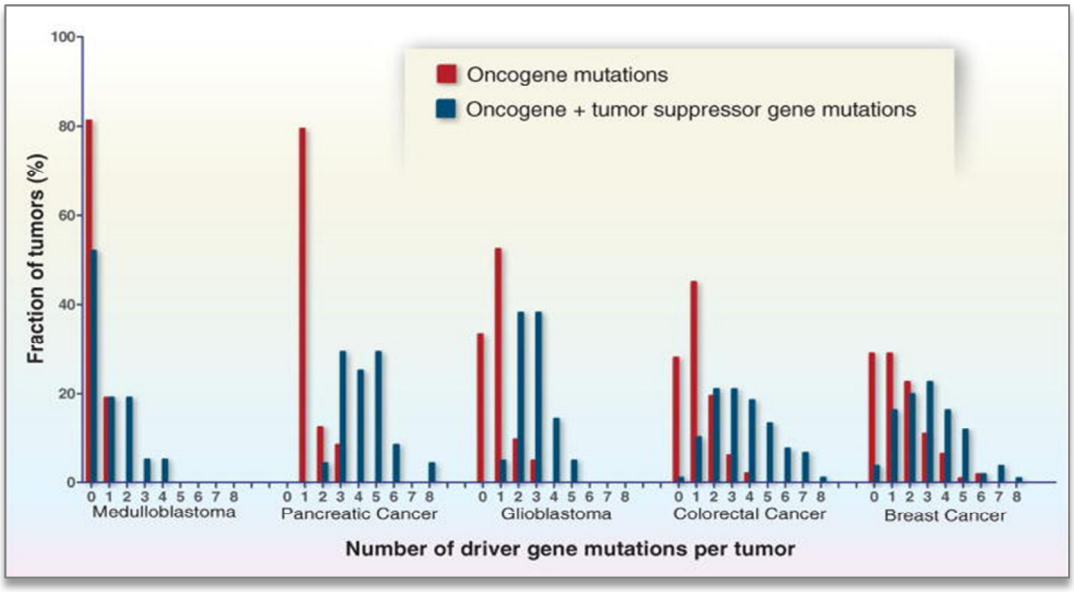


Figure 4: Number of driver gene mutations in five tumour types [Vogelstein et al. (2013)].

Tumours acquire a series of mutations over time and the transformation into a malignant state may need decades. The first mutation increases the growth rate of the cell and occurs frequently in known driver genes. These mutations are called gatekeeping mutations and allow the cell to outgrow the neighbour cells. The first mutations, which triggers a slowly growth, is followed by second mutation, such as in the *KRAS* gene, which admits an expansion of cell number. Additional mutations in genes like *PIK3CA* and *TP53* form finally a malignant tumour that can invade the basal membrane and metastasize to distant tissue. The number of mutations is correlated with age in certain self-renewing tissues explaining that adult tumours have more mutations than paediatric cancers as they often occur in non-self-renewing tissues. It helps to explain why brain tumours and pancreatic tumours show fewer mutations than breast cancer; brain cells and pancreatic cells usually do not replicate and belong therefore to the non-self-renewing tissue [Vogelstein *et al.* (2013)].

Furthermore, most tumours display changes in chromosome number and other genetic abnormalities such as deletions, inversions, and translocations. Two genes can be attached to each other to create an oncogene due to translocation. Translocation can also shorten genes and therefore inactivate tumour suppressor genes. However, the majority of translocations seem to be passengers rather than drivers. Deletions occur just in few tumour suppressor genes, whereas amplifications involve an oncogene with an abnormally active protein product [Vogelstein *et al.* (2013)].

Cell division is a physiological process in normal cells including mutations of genes that does not affect the viability. Recent studies estimated up to 2,500 structural variations such as duplication, insertions, and inversions and approximately three million single nucleotide polymorphisms between any humans [O'Huallachain *et al.* (2012)]. As mentioned above, variations between tissues within one individual have been revealed in the past. It seems likely that somatic mutations present a growth advantage to the affected cells [O'Huallachain *et al.* (2012)]. Intratumour heterogeneity therefore might foster tumour evolution and adaptation. Gerlinger *et al.* (2012) uncovered branched evolutionary tumour growth in which 63 - 69 % of all somatic mutations cannot be detected within the primary carcinoma and associated metastatic sites. They proved intratumour heterogeneity for several tumour suppressor genes as well as intermetastatic heterogeneity. During tumour progression intratumour heterogeneity arises due to the evolution of genetically diverse subclones. Metastatic sites show not only heterogeneity among different locations, but as the metastases grow heterogeneity develops among the cells of each metastasis [Vogelstein *et al.* (2013)].

Genetic heterogeneity among the cells of an individual tumour exists constantly and challenges the development of biomarker and personalized medicine strategies. Genetic dynamics and single-cell heterogeneity has an impact on the response to therapeutics, therefore targeted therapies directed at specific molecular states lead to better treatment of cancer patients compared with common therapy and need further conceptual and technological improvements [Beckman *et al.* (2012)].

1.2.2 Chromosomal aberration- Aneuploidy

The point mutation rate in tumours equals more or less the rate of normal cells. In addition to activated oncogenes and mutation in tumour suppressor genes most tumours frequently contain rearrangement of the chromosomes [Swarup, Rajeswari (2007)].

One of the characteristic of the cancer cell genome is numerical and structural alteration. Aneuploidy, one feature of human malignancies refers to a cell state, in which number of chromosomes is abnormal and is not an exact multiple of two. It was first associated with cancer one century ago, but its impact in cancer is still a subject of debate. Some state that aneuploidy is a benign side effect of cellular transformation, whereas others have considered aneuploidy as a core element contributing to the growth, development and adaptability of tumours. Approximately 90 % of solid human cancer cells show an abnormal number of chromosomes with varying degree and spectrum among the tumour types [Holland, Cleveland (2012)]. Tumour cells can have not only numerical alterations but structural alterations of chromosomes, including deletions, duplications, inversions, and translocations.

McGranahan *et al.* (2012) describes chromosomal instability (CIN) as "a dynamic state in which cells continuously gain or lose whole chromosomes, or parts of chromosomes, at an elevated rate". CIN is therefore an active process that leads to the state of aneuploidy and also intro-tumour heterogeneity. In clinic, CIN is linked to poor patient outcome in a variety of cancer types, including breast, colon, and lung cancer [McGranahan *et al.* (2012)]. It shows a subsetting into two groups. First, numerical chromosome changes are caused by the inability of accurately segregate whole chromosomes. Structurally abnormal chromosomes on the other hand results in gain or loss of chromosome fragments, translocations, deletions and amplifications of DNA fragments (Figure 5). Structural rearrangements can alter genes including formation of fusion gene products and amplification, or deletion of some genes. CIN leads to aneuploidy all the time [McGranahan *et al.* (2012)], whereas aneuploidy does not necessarily involve CIN as cells can have a stable abnormal number of chromosomes. Chromosomal segregation errors in cell division causes aneuploidy, while structural chromosomal alterations are the result of wrong repair of DNA double stand breaks [Ricke, van Deursen, J. M. (2013)].

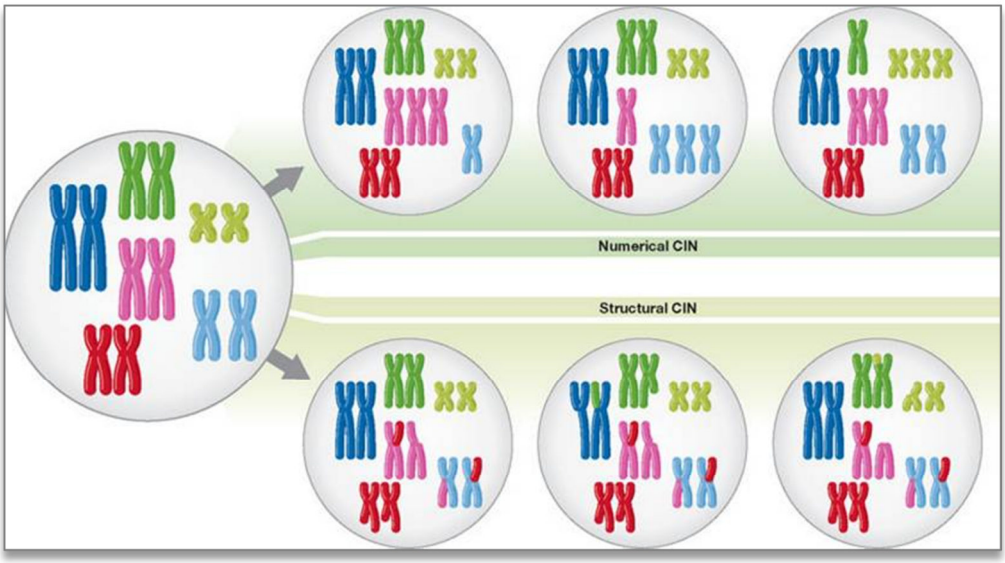


Figure 5: Numerical and structural chromosomal instability (CIN) [McGranahan et al. (2012)].

To determine the aberrant number of chromosomes or structural variants a variety of different technologies are on the market such as comparative genomic hybridization (CGH), fluorescence *in situ* hybridization (FISH), flow and DNA image cytometry, and next-generation sequencing systems [McGranahan *et al.* (2012)].

Aneuploidy is the most frequent origin of miscarriage and mental retardation in humans. Most abnormal numbers of chromosomes are lethal; only three autosomal trisomies are viable including trisomy 21 leading to Down syndrome [Holland, Cleveland (2012)]. Furthermore, aneuploidy is not only limited to cancer or other diseases such as Down syndrome. Indeed, an abnormal number of chromosomes is associated with normal organ development and aging as pointed out by Ricke, van Deursen, J. M. (2013). First, aneuploidy

might be beneficial for some tissues like brain and liver. It has been suggested that aneuploidy endorses the diversity of brain cells contributing to plasticity, which is necessary to complex tasks including learning and memory. Cells of the liver have an abnormal number of chromosomes and it is thought that this process provides the cells a selective advantage protecting the liver against unknown attacks. Aneuploidization is also a characteristic of aging, which contributes to tissue degeneration.

The knowledge about relationship and effect of aneuploidy/CIN and cancer is yet very limited. Down syndrome patients for example are less affected by solid tumours and have a significant decrease in haematological cancers located in blood circulation, bone marrow, and lymph nodes [Holland, Cleveland (2012)]. Chromosomal stability is believed to be influenced by hundreds of genes with some being more cancer relevant than others. Ricke, van Deursen, J. M. (2013) classified the effect of CIN gene alterations within cell division (mitosis) regulators into three groups. The first group describes CIN gene defects in human cancers that can supress tumour development. The second group of CIN gene alterations has either little or no impact at all on tumour development. Third, genes of mitotic regulators are altered in cancer, causing aggressive tumour progression.

„Cancer cells affected with CIN missegregate one chromosome every one to five divisions in vitro" [Ricke, van Deursen, J. M. (2013)]. CIN has been always linked to poor prognosis. Furthermore, CIN characteristic results in diversity in tumour cell population that supports the tumour cell responding to changing selection pressure. Therefore, scientists argue that aneuploidy promotes tumour formation. On the other hand, cancer cells need to be capable to respond to changing intracellular and extracellular environment in order to grow continuously. So reducing the proliferation rate might enable the tumour to adapt and progress better and therefore aneuploidy can suppress tumour formation in certain circumstances. Thus far, it is shown that a slow proliferation rate in human colorectal cancer makes the tumour more aggressive and indicates an increased ability to metastasize. Therefore, the effect of aneuploidy on tumour progression depends on the tissue microenvironment, and the genetic context [Holland, Cleveland (2012)]. It is difficult to use the CIN status for prognostic and therapeutic targeting, as CIN has been associated with unfavourable patients outcome and problematic drug resistance on the one hand and improved patient survival on the other [McGranahan *et al.* (2012)].

1.3 CANCER IN CLINIC

Statistically, every year in Austria 37,000 to 38,000 people are newly diagnosed with malignant neoplasms with more than half being diagnosed with lung, breast, prostate or intestinal cancer. The risk of a person suffering from one of those cancers before the age of 75 was around 10 % in the year 2011. This means that approximately 10 people in 100 are diagnosed with one of these four cancers before they turn 75 years old [Statistics Austria (2013)].

For men as for women the malignant neoplasms are the second most common cause of death after cardiovascular diseases. Although the number of newly diagnosed cancer incidences increased within the last two decades, fewer people died because of cancer [Statistics Austria (2013)] and therefore the cancer mortality rate is decreasing. Generally, women have an advantage in cancer survival, but the sex difference was reduced in the last decade as men's survival has increased at a greater speed [Statistics Austria (2013)]. An earlier diagnosis and new possibilities of treatment have led to a longer survival time of the patients. A Eurocare study, whose aim is to provide an updated description of cancer survival time across European countries, claims that survival in Austria was above the European mean [Angelis *et al.* (2013)]. In figure 6 the most common cancer types in Austria for male are displayed. Comparing the graph with figure 7, which shows the most frequently cancer types for Austrian females, it differs in the number of neoplasms as men are affected slightly more often than women.

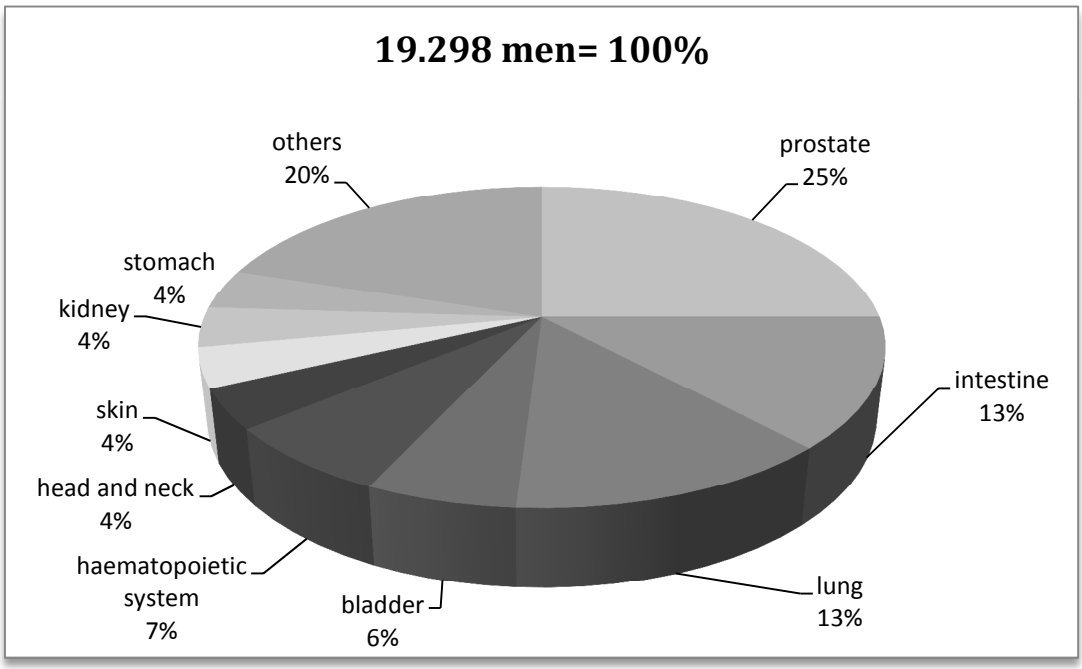


Figure 6: The most common malignant neoplasms in men 2011 [Statistics Austria (2013)] .

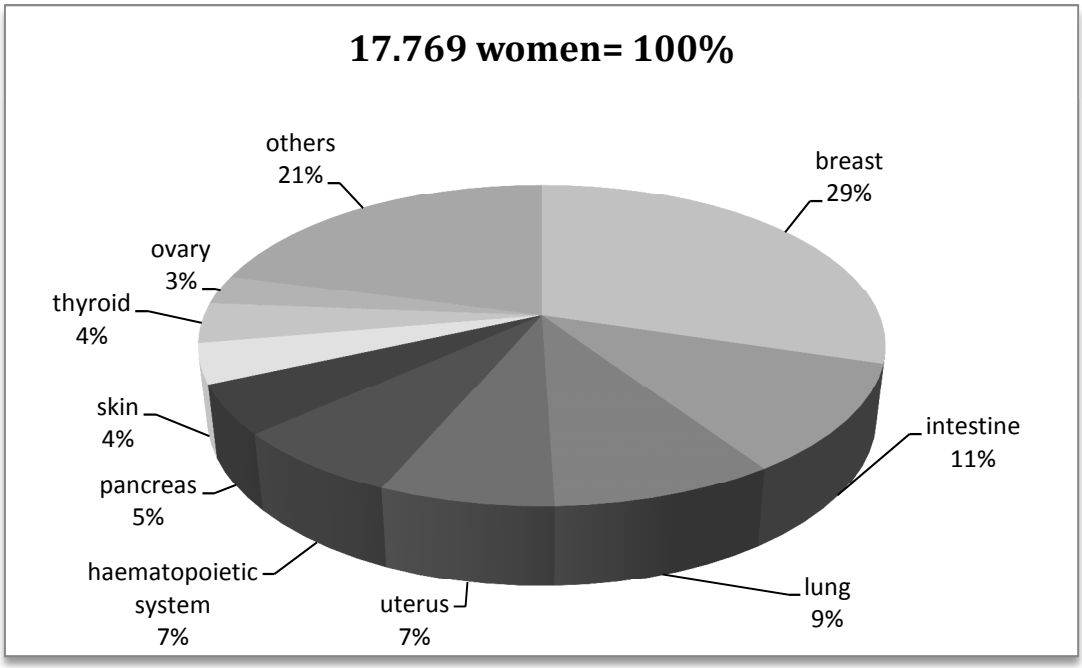


Figure 7: The most common malignant neoplasms in women 2011 [Statistics Austria (2013)].

The stage at cancer diagnosis is an important factor in the prognosis of a tumour and is therefore recorded as far as possible for all cancer incidences for survey of Statistics Austria. Approximately one third of all cancers were diagnosed while the cancer was still limited to the particular organ (localised cancer stage: 32.6 %). One fifth of all new diagnoses in 2011 were only made after the cancer had spread to other parts of the body (regional lymph node metastases, regionalised cancer stage: 18.2 %). Distant metastases were discovered in a further 10.6 % of diagnoses (disseminated cancer stage) [Statistics Austria (2013)].

1.3.1 Prostate cancer

Since 1994, the most frequently diagnosed cancer in Austrian men has been prostate cancer and the incidence rate is currently 24 %. The most common cancer before 1994 was lung cancer that still remains the most common cause of cancer mortality. However, in 2011 4,722 men were newly diagnosed with prostate cancer and 1,146 men died as a result of this type of cancer. Thus, every tenth cancer death was caused by prostate cancer [Statistics Austria (2013)].

1.3.2 Breast cancer

The most common cancer site in women continues to be the breast with 5,434 incidents in 2011. About one woman out of thirteen suffers from breast cancer before the age of 75. According to Statistics Austria, the age-standardised incidence rate of breast cancer decreased by 4 % in the past ten years. The recently diminished incidence of breast as well

as prostate cancer can be attributed to a big part to the so-called screening effect. The increase of the number of preventive medical check-ups and early detection contribute to the increasing incidence of those two frequently diagnosed cancers [Statistics Austria (2013)].

1.3.3 Colon cancer

Colon cancer comprises malignant neoplasms of the colon and the rectum and makes up 13 % of all cancers of men being the third most common cancer. In women, colon cancer is the second most common cancer, responsible for 11 % of all types of cancer in females. Two third of the malignant neoplasms occur in the colon, whereas almost 30 % affect the rectum, the remaining cancers are found in the border crossing between rectum and colon and anal canal [Statistics Austria (2013)].

1.4 DNA analysis

In the area of genetics many new methodologies were developed over the past decades, which enable previously impossible experiments to study in detail the biology of genes. Nowadays different methods exist to measure the more accurate quantity of the DNA and to analyse rapidly the sequences of the nucleotides of the whole human genome. The results obtained from the new techniques allow an understanding how altered gene activity can cause diseases such as cancer.

1.4.1 Polymerase chain reaction

The polymerase chain reaction (PCR) plays an important role in biological research as it is a relatively simple procedure. This revolutionary methodology was invented in the 1980s by Kary Mullis and it extended the range of possible DNA analysis. PCR results in an amplification of a short region of the DNA carried out by an enzyme. The chosen region is amplified in a single test tube by mixing the DNA of interest with a variety of reagents and placing the tube in a thermal cycler. This instrument incubates the mixture at a series of pre-programmed temperatures.

As long as the border regions of the sequence are known, any region of a DNA molecule can be amplified. Those elements have to be known in order to hybridize oligonucleotides, so called primers, to the strand so that the chosen region can be amplified. The basic steps of the PCR are shown in figure 9. The PCR is so simple compared to normal DNA synthesis, because it mimics the cell's own DNA replication process. Few basic building blocks and an enzyme are supplied in a test tube, which allows to specifically target and amplify certain sequences of DNA. Specificity is achieved through the use of the appropriate oligonucleotide

primers. Multiple cycles of this DNA synthesis using the resulting templates significantly amplifies the target DNA sequence. As PCR is extremely sensitive, it can detect even one DNA molecule in the starting mixture. The primers can for example be designed for a clinical application to detect the DNA of a disease causing virus.

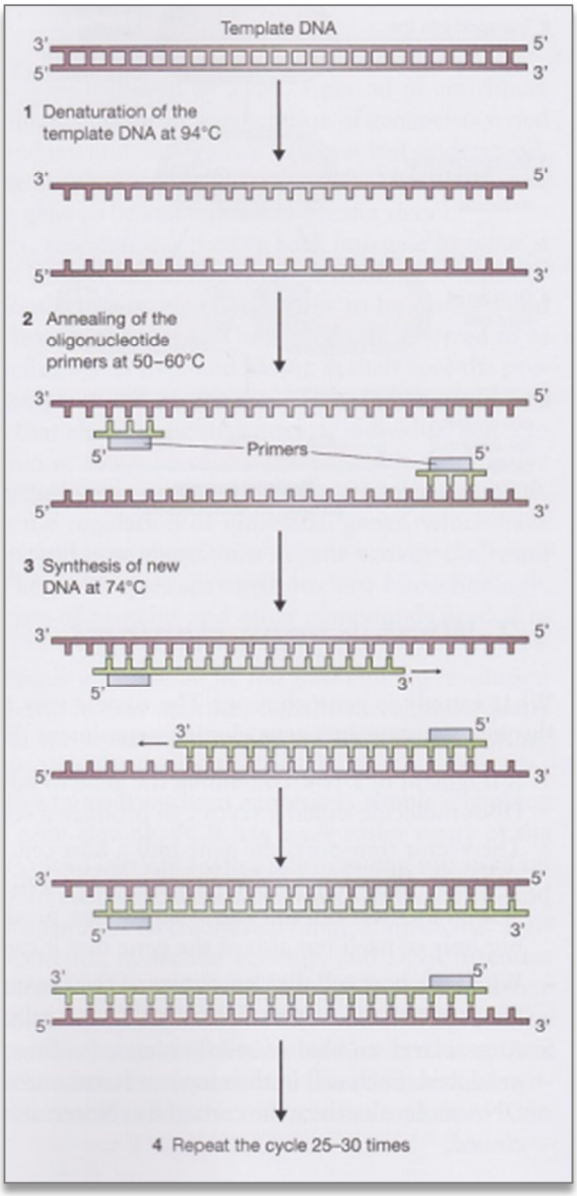


Figure 9: The basic steps of the PCR [Brown (2010)].

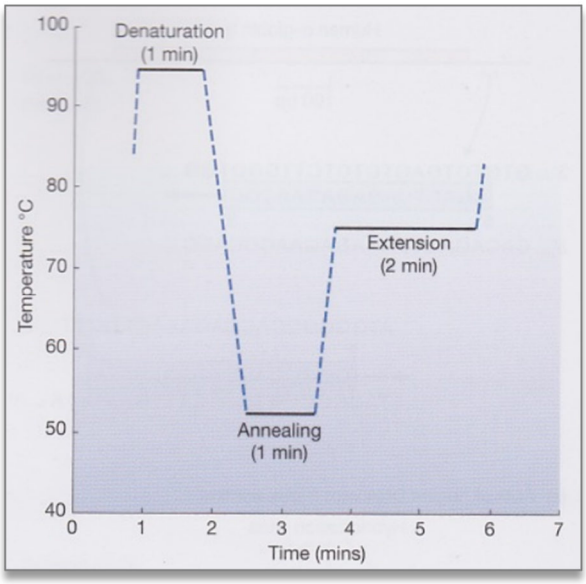


Figure 8: Temperature profile for a PCR [Brown (2010)].

First, the mixture of DNA, primers, nucleotides, and enzymes is heated above 90 °C to break the hydrogen bonds of the template double stranded DNA (dsDNA) helix causing it to denature and become single stranded (ssDNA). The sample is cooled down to 50 °C to 60 °C to allow oligonucleotide primers to anneal to the ssDNA at specific regions. A big excess of oligonucleotide primers prevent the single strands of DNA to bind to each other again, what they would do in this temperature range. Each primer must be complementary to the sequences flanking the target region on the template and the 3'ends of the primers need to

point toward one another. The temperature is raised then to the working temperature of 74 °C for a heatstable DNA polymerase such as derived from the bacteria *thermus aquaticus*. This enzyme attaches to the end of the primers and synthesizes the complementary second strand using the DNA strand as a template.

To begin the second cycle of amplification, the temperature is increased again above the 90°C to denature the dsDNA molecules that consists of one original strand and one new strand of DNA. The cycle of denaturation, annealing and synthesis is repeated about 25 to 30 times which results in an amplification of the specific region to millions of copies. The short DNA sequence is doubled during each cycle and is accumulated in an exponential fashion until the PCR reagents (nucleotides, primers or DNA polymerase) are depleted.

As shown in figure 8, PCR works with three different temperatures: the denaturation temperature to break the bonds of the dsDNA, the annealing temperature that allows the primer to attach to the templates and the extension temperature, at which the DNA polymerase synthesise the second complementary strand. All three temperatures are very crucial and choosing the wrong settings results in no amplifications, as denaturation, annealing and extension are not carried out in a proper way.

The PCR product is usually analysed later on to check if sufficient copies of the target DNA has been produced. Therefore, gel electrophoresis, cloning and sequencing are three important techniques to study the PCR products. The third technique is dealt with in the chapter 1.4.3. Typically, the PCR is used at the beginning of a study where its product is examined to gain information about the DNA sequence that can for example code for an oncogene.

Although the PCR amplification is completed within few hours it also has some limitations. First, the sequences at the border of the region of interest must be known in order to anneal the primers. Therefore, PCR cannot be used to isolate and amplify unknown genes that have not been studied before. To overcome this limitation the equivalent gene of a different organism could be used to design the primers for the human gene. Second, the length of the DNA molecule that can be amplified by PCR is limited. The ideal length of the DNA fragment that should be amplified is less than 1 kb, but fragments up to 10 kb can be amplified. Most human genes however, are too long to copy it with specialised PCR techniques, which have the capability to amplify sequences up to 40 kb in length. Therefore, other more complex methods as gene cloning are needed [Brown (2010)].

1.4.2 REAL TIME PCR

The PCR can also be used to quantify the amount of starting DNA. The amount of final product amplified by PCR depends on the number of DNA molecules in the starting mixture, the more DNA helixes are present in the starting mixture the greater the product yield will be. This can be measured through fluorescence.

In order to amplify and quantify fragments of the DNA molecule the same method is used as in conventional PCR, real time PCR monitors the progress of the PCR as it occurs and therefore the data is collected throughout the amplification rather than at the end of the PCR (Figure 10). The amount of fragments synthesised during that PCR run is compared with known quantities of starting DNA that are amplified in parallel.

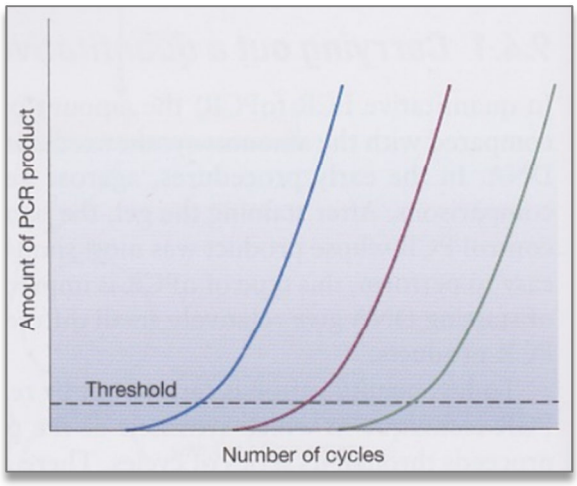


Figure 10: Quantification by real time PCR. Three different PCR products with a different amount of starting DNA that are quantified through a threshold [Brown (2010)].

Quantitative PCR is carried out on a thermal cycler, which is able to detect the fluorescence emitted by a fluorochrome while heating and cooling the samples. Two types of qPCR techniques exit using different dye reporter:

(1) First, a dye, which emits fluorescence signals when bound to the dsDNA, can be added to the PCR mixture at the start. The polymerase synthesises more PCR products in every cycle and the dye binds to every new copy of the dsDNA. The fluorescence intensity is proportional to the amount of PCR product. Therefore, the more the dsDNA product is increased during the PCR the more increases the fluorescence intensity that is measured by a detector. However, the dye such as SYBR® green binds to all double stranded DNA molecules during the PCR. It can therefore overestimate the amount of PCR product, because it can also bind to non-specific dsDNA sequences such as primers annealing to each other to lead to an increased amount of dsDNA. Another aspect of using this method is that the intensity of the fluorescence might depend on the mass of the dsDNA and an amplification of a longer fragment will emit more signal than for a shorter one.

(2) The second method uses a fluorogenic-labelled oligonucleotide called a reporter probe (Figure 11) that emits a fluorescent signal when hybridised to the PCR product. The probe allows a specific hybridisation to the target sequence of the PCR product and increases specificity without for example the detection of primer-primer annealing. The probe has at one end the fluorescent reporter and on the opposite end the quencher of fluorescence attached. The quencher inhibits the emitting of the fluorescence by absorbing the energy when it is close enough to the dye. Once the probe hybridises to the DNA target the hair pin structure of the probe, which brings the dye close to the quencher, is disbanded. Thus, the quencher separates from the dye, which enables the detection of the fluorescence signal.

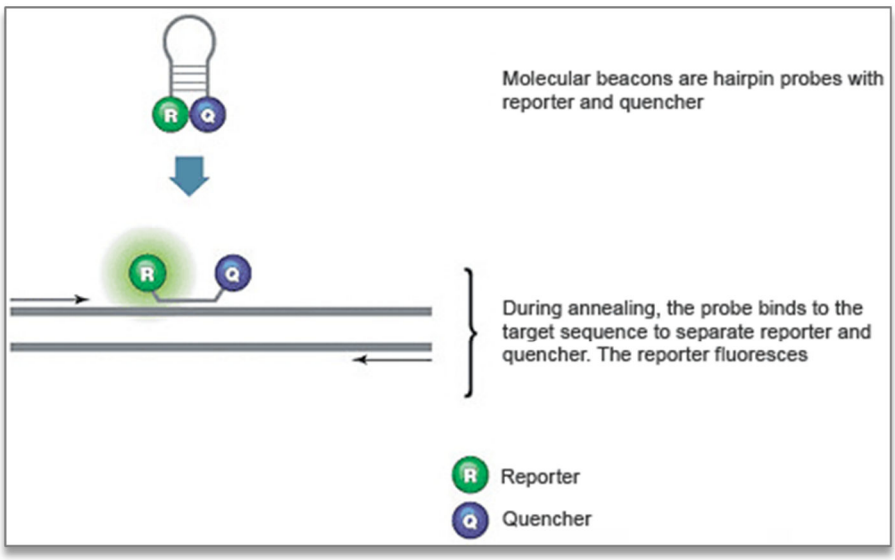


Figure 11: Reporter probe is hybridised to target DNA [BIO-RAD (2014)].

Both methods measure the increase of the fluorescence depending of the higher yield of PCR products. To be able to quantify the test DNA the fluorescence intensity is compared to the intensity of a control DNA of known starting amount. Therefore, the cycle at which the amount of PCR product moves about a pre-set threshold is identified for the comparison (Figure 10). The earlier the threshold is reached during the PCR, the bigger is the amount of starting DNA in the test tube.

1.4.3 NEXT GENERATION SEQUENCING

DNA fragments are sequenced in order to obtain the order of nucleotides of individual genes or entire genomes. Since the beginning of sequencing in the 1970s novel technologies has been developed and the process of sequencing became cheaper and much faster in the past decade. At the beginning, sequencing of even individual genes required a lot of time, whereas today the entire bacterial genome of *Escherichia coli* can be determined in less than one day using automated and massive parallel sequencing technologies [Loman *et al.* (2012)]. From the time of invention onward, research accelerated due to innovations and automation. Different methods of high-throughput sequencing platforms have been published to date.

The chain-termination method, also called Sanger sequencing named after its developer, is the method of choice to sequence the desired order of nucleotides. Sanger Sequencing is the most popular procedure and was used for example to decode the human genome. This method is using oligonucleotide primers, DNA polymerase and in addition to the normal nucleotides different fluorescently labelled chain-termination nucleotides, whose fluorescent signals are distinguished as they pass by a detector. Cycles of denaturation, primer annealing and strand extension are performed, whereas the extension process is stopped once one of the labelled dideoxynucleotides is incorporated to the newly synthesised strand by the DNA polymerase [Sanger *et al.* (1977)].

A novel strategy was to design a faster system that enables a large number of sequencing experiments at the same time and the use of new methods such as pyrosequencing. In the last few years, next-generation sequencing (NGS) technologies have increased the speed of DNA sequencing compared to Sanger sequencing [Zong *et al.* (2012)]. NGS enables high-throughput DNA sequencing and is applicable to very large scale applications such as whole genome studies or studies about genetic and epigenetic differences between cells. NGS employs micro- and nanotechnologies that reduce the size of sample components, reagent costs and enables simultaneous sequencing and analysis of millions of samples. Therefore, applications of NGS are not limited to the detection of DNA sequence variations associated with inheritable human disorders and whole-genome sequencing for reference genomes, but also include diagnosis, prognosis and treatment of cancer and infectious- disease testing in laboratory medicine.

The most widely used platforms for whole-genome sequencing are sold by Illumina and Complete Genomics [Lam *et al*. (2012)]. Besides these two there are many more platforms with different NGS technologies [Einstein (2012)]. The fast moving field is steadily improved by new technologies and improvements in quality. The first NGS technology on the market was offered by Roche and is based of emulsion PCR and pyrosequencing. Illumina released in 2006 sequencing-by-synthesis technology that uses reversible fluorescent dye terminators.

Later on, Life Technology developed the Ion Torrent platform, which is based on semiconductor chips with sensors that detect protons liberated in the course of nucleotide addition. Those instruments differ in their performance with respect to sequence coverage (completeness of sequencing), insertions and deletions [Lam *et al.* (2012)] as well as in cost per run, run time, minimum throughput, and the read length. Figure 12 represents the properties and mechanism of selected sequencing technologies such as SOLiD/Ion Torrent PGM from Life Sciences, Genome Analyzer/HiSeq 2000/MiSeq from Illumina, and GS FLX Titanium/GS Junior from Roche.

Sequencer	454 GS FLX	HiSeq 2000	SOLiDv4	Sanger 3730xl
Sequencing mechanism	Pyrosequencing	Sequencing by synthesis	Ligation and two-base coding	Dideoxy chain termination
Read length	700 bp	50SE, 50PE, 101PE	50 + 35 bp or 50 + 50 bp	400 ~ 900 bp
Accuracy	99.9%*	98%, (100PE)	99.94% *raw data	99.999%
Reads	1 M	3 G	1200~1400 M	–
Output data/run	0.7 Gb	600 Gb	120 Gb	1.9~84 Kb
Time/run	24 Hours	3~10 Days	7 Days for SE 14 Days for PE	20 Mins~3 Hours
Advantage	Read length, fast	High throughput	Accuracy	High quality, long read length
Disadvantage	Error rate with polybase more than 6, high cost, low throughput	Short read assembly	Short read assembly	High cost low throughput

Figure 12: Methodology and properties of different NGS instruments [Liu et al. (2012)].

Although the technologies used by the platforms differ from each other, the methods follow the same principal workflow. In order to achieve identical DNA fragments as the sequencing material, the DNA of interest is fragmented and amplified by PCR attaching specific adapters for the used platform. After the library preparation step, the double stranded PCR products are denatured into single strand molecules and sequenced using one of the NGS technologies.

NGS has a much higher degree of parallelism and needs less hands-on steps than Sanger sequencing, but it also needs more effort to create the indexed fragmented libraries, and the handling of the huge amount of data still remains a challenge.

Furthermore, it is difficult to achieve a gapless, errorless and end-to-end assembly of a complete genome. Repetitive regions like satellite sequences, ribosomal sequences, or transposons still remain a challenge to sequence even in small genomes. Thus, *de novo* genome assemblies by NGS tend to fail, because identical reads might be generated from multiple locations. Further improvements to create longer read lengths or including additional information might overcome this problem. Practically, Sanger sequencing is still used to validate mutations due to its hierarchical property that allow more accurate sequencing results [Gerlinger *et al.* (2012)]. However, according to the publication from Sikkema-Raddatz *et al.* (2013), targeted NGS can replace Sanger sequencing in clinical diagnostics.

1.4.4 ILLUMINA SEQUENCING- MISEQ

One of the NGS instruments that are available to date is the MiSeq (Illumina) capable of sequencing an entire genome within a day. Illumina sequencing is based on cluster generation and sequencing-by-synthesis technology and enables an accurate and fast sequencing on a large scale. First the genomic DNA samples are prepared, for that the DNA is randomly fragmented and adapters are ligated to both ends of the fragments. The single strand DNA fragments are immobilised on the surface of the flow cell for the cluster generation by synthesis of DNA (Figure 13). Adding nucleotides and enzymes initiates the solid-phase amplification that creates up to 100 copies of each DNA template using the principle of PCR. The enzyme adds nucleotides to the single strand to build a double-strand bridge on the flow cell.

The double-stranded molecule is denatured resulting in single-stranded templates anchored to the surface. A result of this repeated bridge amplifications are million dense clusters of double stranded DNA in every channel of the flow cell generated by the first synthesis step. The previous cluster generation follows a second synthesis of DNA step to identify the sequence of the nucleotides in the clusters. Yet, this sequencing-by-synthesis (SBS) technology uses fluorescently labelled nucleotides to read out the sequences of the cluster base pairs. For the first sequencing cycle primer, DNA polymerase and all four labelled nucleotides are added to the flow cell. Corresponding to the template a single nucleotide is incorporated to the chain. Due to its label the nucleotide terminates the polymerisation, so that after every incorporation of a nucleotide the fluorescent signal is imaged to identify the base.

After laser excitation, the emitted signal is captured and the label is cleaved by an enzyme to make the next incorporation in the chain possible. For the next cycle primers, DNA polymerase and the four reversible labelled nucleotides are added and the fluorescent signal is captured as previous. Those cycles, one base at a time, are repeated to define the entire sequence of the nucleotides in the fragments. This automated workflow requires barely hands-on time and sequences quickly several gigabases of DNA per run. The Illumina sequencing technology of Illumina usually produces 50 – 200 million 32 – 100 bp long reads on a single run of the machine [Li, Durbin (2009)].

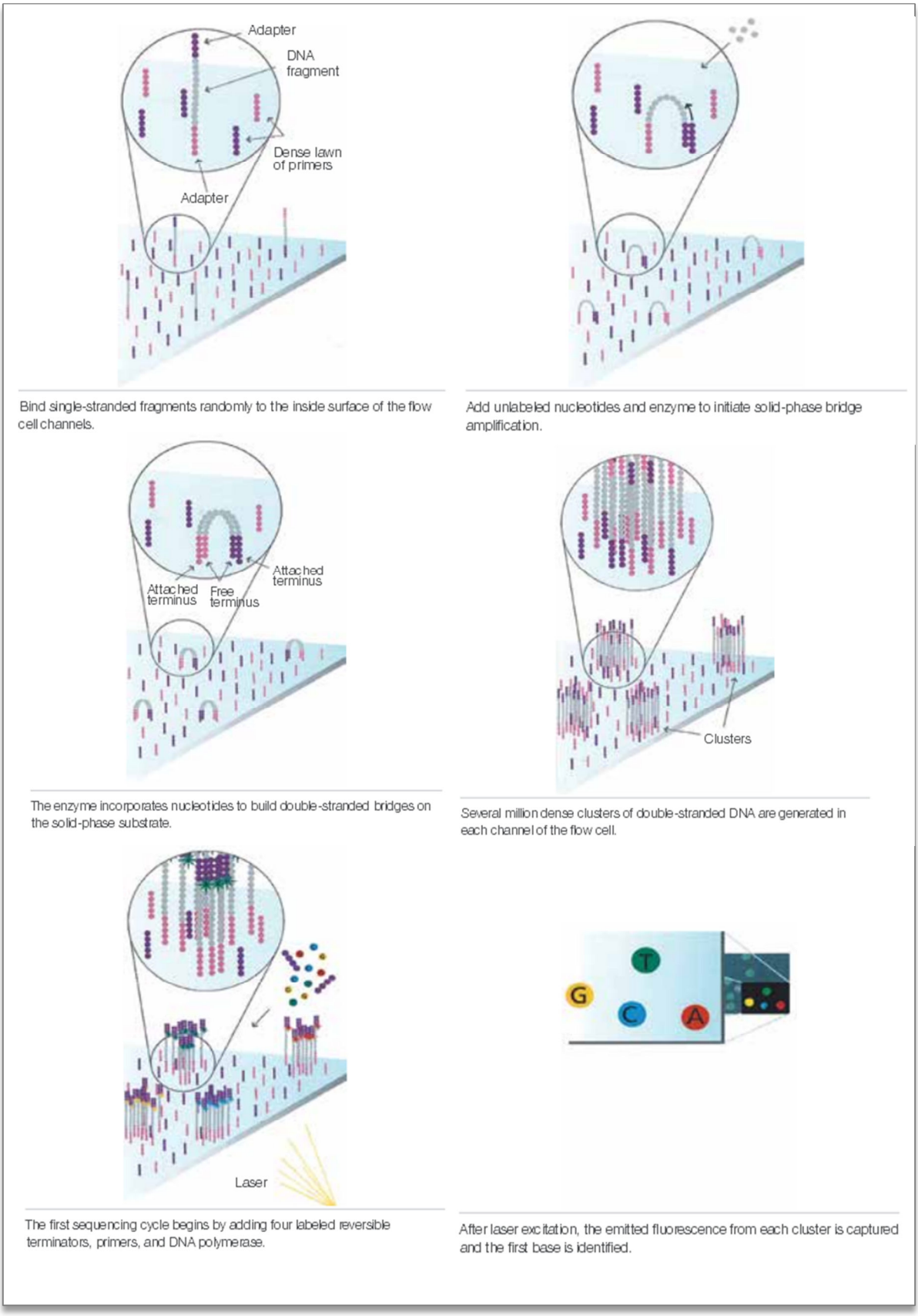


Figure 13: Illumina Sequencing Technology; Cluster generation and sequencing by synthesis [Illumina (2010)].

1.4.5 PLASMA DNA FOR NON-INVASIVE IDENTIFICATION OF CANCER

Lo *et al.* (1998) developed a real time PCR method to measure the concentration of fetal DNA in the maternal plasma and serum. They showed that the fetal DNA is present in high concentration in the maternal plasma [Lo *et al.* (1998)] and increases as pregnancy progresses. From this point on, researchers investigated the maternal plasma and serum as a valuable source for non-invasive prenatal diagnosis [Kinde *et al.* (2012); Forshew *et al.* (2012); Fan *et al.*(2012)]. Circulating cell- free DNA existing during pregnancy was first discovered in 1948. Later on, many studies also identified cell- free circulating DNA extracted from cancer patients [Chan, K. C. A. *et al.* (2013)] that originate from the tumour itself. Today it is well known that the release of small amounts of DNA into the plasma occurs also in healthy individuals [Swarup, Rajeswari (2007)].

Circulating cell free DNA (cfDNA) are small fragments of genomic DNA present in the plasma or serum. Circulating DNA has a length of 140 to 170 bp with a frequency of a few thousand amplifiable copies per millilitre blood [Forshew *et al*. (2012)]. A study by Dawson *et al.* (2013) showed, that circulating tumour DNA shows even a greater correlation with tumour changes than circulating tumour cells, which originate from solid tumours.

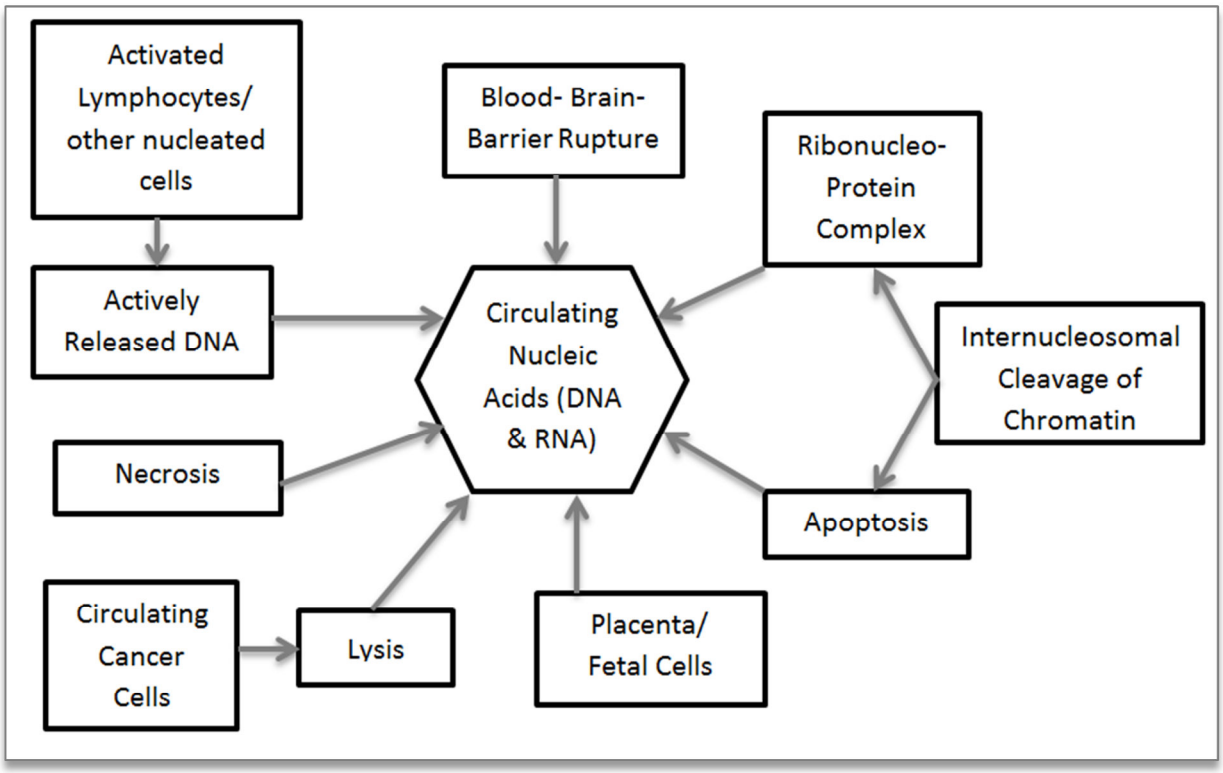


Figure 14: Various pathways by which DNA can be released in the plasma [Adopted from Swarup, Rajeswari (2007)].

The exact origin of the circulating DNA and RNA in plasma remains unclear but as shown in figure 14, several possible pathways have been formulated. In healthy individuals, DNA could be released in the circulation via apoptosis. Furthermore, the unstable RNA perhaps binds to protein or phospholipids and is so protected from degradation. In cancer patients, lysis, necrosis, apoptosis, and active release of the DNA in the plasma are possible pathways for the circulating DNA. Three of those cases cause cell death due to decomposition and break down of tumour cells. Lymphocytes have demonstrated the capacity to release DNA in the absence of any stimulation [Anker *et al.* (1976)]. This process is unrelated to cell death and differs from the inactive DNA release caused by cell lysis or necrosis.

Other pathological disorders such as stroke, autoimmune disorders, myocardial infarction, diabetes, trauma, and prion diseases show a relationship between injured tissue and an elevated plasma DNA level. Rainer (2003) showed in his studies that plasma DNA concentrations correlate with stroke severity and can therefore help to predict mortality and morbidity of acute stroke patients. Based on the study of Chang *et al*. (2003), it appears that the plasma DNA concentration in myocardial infarction patients is 10 times higher than in the controls, proving a direct relation between the extent of injury and the increase of circulating DNA in plasma.

Many scientist have shown that the increased levels of circulating plasma DNA as well as plasma RNA can be used as a non-invasive, rapid, sensitive and accurate method of diagnosis of several diseases [Swarup, Rajeswari (2007); Kinde *et al.* (2012); Leary *et al.* (2012); Forshew *et al.* (2012); Beckman *et al.* (2012)].

In addition to cancer screening, parallel sequencing of cell-free, maternal plasma DNA can be also used as a safe and effective screening method for fetal chromosomal aneuploidies [Kinde *et al.* (2012)]. According to Kinde *et al.* with their improved sequencing method, namely FAST-SeqS, samples containing as little as 4 % of trisomy 21 DNA, causing Down syndrome, could be detected. There is no technical barrier anymore to determine the entire fetal genome through the maternal plasma by non-invasive methods [Fan *et al.* (2012)].

Cancers are characterised by heterogeneity and genetically dynamic state. Non-invasive methods could provide real-time information about the dynamic nature of cancer. Through the development of non-invasive methods for detection of tumours clinical management of the cancer patients can be improved, because conventional methods of cancer screening are often invasive and expensive. Using plasma DNA for cancer screening is an approach for such non-invasive method since the cell-free circulating DNA levels are high in patients with malignancies [Leary *et al.* (2012)]. Losses and gains of chromosomes, alteration of chromosome arms, focal amplifications and deletions, and rearrangements are found in almost all malignant genomes such as lung, breast, colon, ovarian, and prostate cancer (Figure 15). Analysis of such chromosomal rearrangements by NGS methods allows a high resolution mapping to evaluate possible tumour specific biomarkers. Those biomarkers can

help in prognosis, diagnosis and in drug design as a predictive biomarker [Ghorbian, Ardekani (2012)]. Quantification of the cfDNA is possible with real-time PCR and other DNA assays.

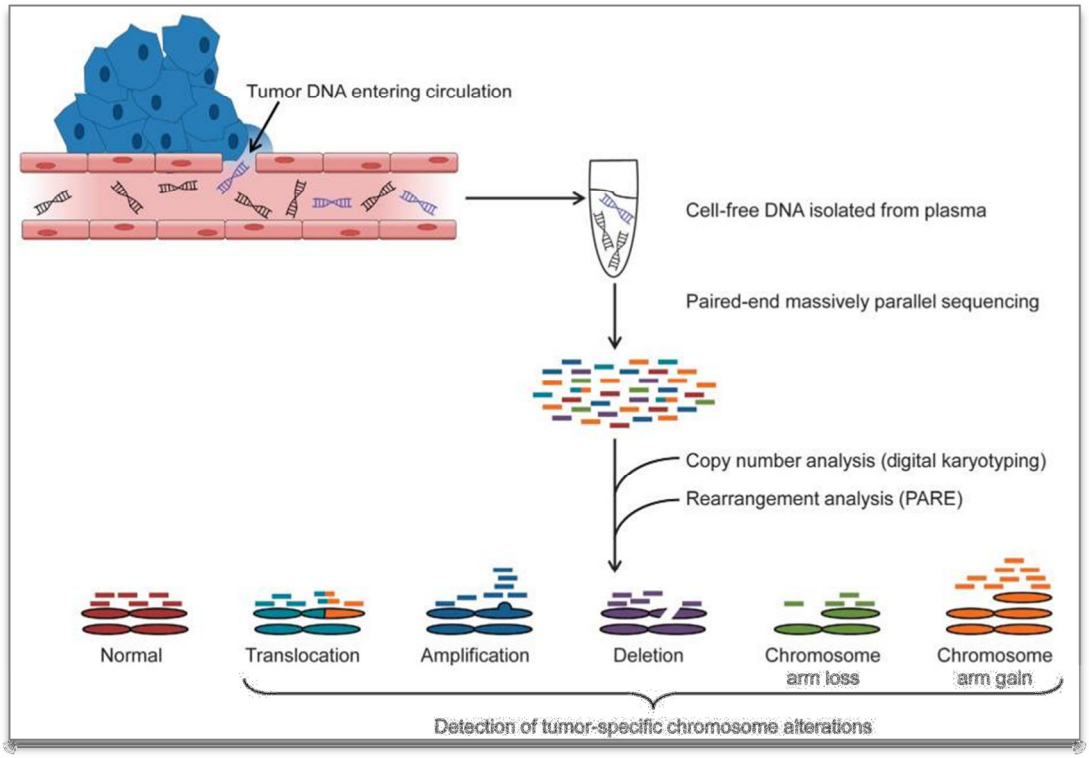


Figure 15: Graphic of analysis for chromosomal alterations in plasma [Leary et al. (2012)].

The average cfDNA concentration of cancer patients is higher than in healthy individuals. The reason for the higher concentration is that the tumour DNA and the non-tumour DNA contribute to the cfDNA level of a cancer patient. However, patients suffering from an inflammation, trauma, premalignant states, and acute or chronic illnesses also have an increased cfDNA level. This observation of high variability can prevent the cfDNA from becoming more than a critical biomarker of cancer. Highly sensitive methods are needed to detect the tumour DNA among the non-tumour DNA in the plasma. However, tumour-specific genetic markers in circulating DNA with high sensitivity and specificity can already be used to detect for example mutations in *TP53*, *NRAS* and *KRAS* genes [Kumar (2006)].

2. THESIS OBJECTIVE

Cancer is the second leading cause of mortality after cardiovascular diseases. Besides mutation of the genome, copy number changes (CNV) of chromosomes are another feature in tumours of different types. Many studies have shown that the increased levels of circulating plasma DNA derived from the tumour can be used for non-invasive cancer diagnosis and monitoring.

Kinde *et al.* (2012) published recently an improved sequencing method, which they named Fast Aneuploidy Screening Test-Sequencing System (FAST-SeqS). This method was originally developed for a faster and cheaper prenatal testing of aneuploidy by massively parallel sequencing of maternal plasma DNA. Instead of a whole genome library preparation and the complex analysis, the preparation process was simplified using only a defined number of repetitive LINE sequences using a single primer pair. This new method may be also applicable for aneuploidy screening of cancer samples.

Hence, the objective was the implementation of this new method for analysis of chromosomal copy number variations in cancer cells. First the sensitivity of detection was evaluated using DNA serial dilutions of four different tumour cell lines. To simulate plasma DNA, the DNA of those four cell lines were fragmented and their dilutions series were also assessed regarding the detection limit. Furthermore, the method was used to analyse the plasma DNA of tumour patients including breast, prostate and colon cancer patients that already have a known CNV profile. Control samples for the tumour sample as well as for the cell line samples were processed and compared to the previous results. The tumour samples were evaluated with plasma DNA samples of healthy individuals that were treated with the same new sequencing method, whereas the results of the dilutions series were compared to the CNV profiles obtained with an established library preparation kit sold by Illumina.

The hypothesis is that the chromosomal copy number variations of tumour samples can be detected with the new screening method that analyses a defined number of LINE sequences. This method is also applicable to non-invasively collected samples such as plasma DNA. Fulfilling those requirements, the detection should be carried out in a fast and cheap way with less analysis time, but higher throughput.

3. MATERIALS AND METHODS

3.1 MATERIALS

In table 1 all instruments that were used are listed and in table 2 common materials are presented which were utilised during the experiments. All specific materials for the single steps of the experiments are listed separately in the next tables.

Table 1: List of instruments used for the study

Instrument	Company
Centrifuge 5417 R	Eppendorf
Shaker MS3 basic	IKA®
Thermomixer compact	Eppendorf
Allegra X-12R Centrifuge	Beckman Coulter
QIAcube®	QIAGEN
Qubit® 2.0 Flourometer	Life Technologies
S220 focused-ultrasonicator	Covaris
SonoLab™ Software	Covaris
2100 Bioanalyzer	Agilent Technologies
2100 Expert software	Agilent Technologies
Bioanalyzer chip vortexer MS3	IKA®
Chip priming station	Agilent Technologies
Step One Plus real time PCR system	AB, Life Technologies
StepOne™ software v2.2.2	AB, Life Technologies
DNA Engine Dyad® Thermal Cycler	BIO-RAD
Dynal MPC®-M Magnetic Particle Concentrator	Life Technologies
Concentrator Plus	Eppendorf
MiSeq Benchtop Sequencer	Illumina
MiSeq Control Software v2.0	Illumina

Table 2: List of standard laboratory equipments

Materials	Company
15 ml tubes	Greiner Bio-One
Safe-lock tubes 1,5 ml	Eppendorf
PCR soft tubes 0,2 ml	Biozym Scientific

Safe-seal tips professional 10 µl	Biozym Scientific
Safe-seal tips professional 20 µl	Biozym Scientific
Safe-seal tips professional 100 µl	Biozym Scientific
Safe-seal tips professional 200 µl	Biozym Scientific
Safe-seal tips professional 1000 µl	Biozym Scientific
Pipetman neo P2N	Gilson
Pipetman neo P10N	Gilson
Pipetman neo P20N	Gilson
Pipetman neo P100N	Gilson
Pipetman neo P200N	Gilson
Pipetman neo P1000N	Gilson

Table 3: List of materials required for cell line harvesting

Consumables	**Company**
QIAamp DNA Micro Kit	QIAGEN
100% ethanol	Emsure

Table 4: List of materials required for the plasma DNA extraction

Consumables	**Company**
Rotor adapters	QIAGEN
Sample tubes	QIAGEN
1000 µl tips	QIAGEN
Proteinase K	QIAGEN
Reagents: AL, AW1, AW2	QIAGEN
100% ethanol	Emsure
Nuclease free water	Promega, Madison

Table 5: List of materials required for the quantification of DNA

Consumables	**Company**
Qubit® dsDNA HS Assay Kit	Life Technologies
Qubit® dsDNA BR Assay Kit	Life Technologies
500 µl thin-walled PCR tubes	Life Technologies
Agilent DNA 7500 Kit	Agilent Technologies

Table 6: List of materials required for the qPCR

Consumables	Company
Fast optical 96 Well Reaction Plate with Barcode 0,1 ml MicroAmp™	Life Technologies
Optical adhesive film Mirco Amp™	Life Technologies
Fast SYBR® Green	Life Technologies
R1-forA and R1-Rev Primer[1]	Microsynth
Standard human genomic DNA	Promega, Madison
Nuclease free water	Promega, Madison

Table 7: List of materials required for the fragmentation of the DNA

Consumables	Company
1×TE buffer	
microTUBE	Covaris

Table 8: List of materials required for the fast library preparation

Consumables	Company
Nuclease free water	Promega, Madison
5× Phusion HF Buffer	ThermoScientific
10 µM forward and reverse Primer (R1-for a and R1-Rev)[1]	Microsynth
10 mM dNTPs	Invitrogen
2U Phusion Hot Start II Polymerase	ThermoScientific
70 % ethanol	Emsure
AMPure XP beads	Beckman Coulter
1× TE Buffer	
10 µM forward Primer (R2-For)[1]	MicroSynth
10 µM reverse Primer (R2-RevA)[1]	MicroSynth

[1] The sequence of the nucleotides are added in the appendices

Table 9: List of materials required for the library preparation by Illumina

Consumables	Company
AMPure XP beads	Beckman Coulter
80% ethanol	Emsure
TruSeq DNA Sample Prep Kit:	Illumina
End repair mix	
Resuspension buffer	
A-tailing mix	
Ligation mix	
DNA adaptor indexes	
Stop ligation buffer	
PCR master mix	
PCR primer cocktail	

Table 10: List of materials required for sequencing

Consumables	Company
MiSeq Reagent Kit v3, 150 cylces	Illumina
HT 1	Illumina
Reagent cartridge	Illumina
PR2 bottle	Illumina
Stock 1,0 N NaOH	Sigma
Laboratory- grade water	
PhiX Control v3	Illumina

3.2 SAMPLES AND WORKFLOW

A variety of genomic and plasma DNA was used for the experiments. The following samples were analysed during this study:

- Genomic DNA of cancer cell lines HCT 116, HepG2, U2OS and MCF7
- Fragmented genomic DNA of cancer cell lines HCT 116, HepG2, U2OS and MCF7
- Plasma DNA of breast cancer patients
- Plasma DNA of prostate cancer patients
- Plasma DNA of one colon cancer patient
- Plasma DNA of control samples (healthy patients)

The plasma DNA was collected from cancer patients for previous studies at the Institute of Human Genetics, whereas the cell lines are cultured permanently at the institute.

For a better understanding for the preparation steps of the DNA, an overview of the workflow of the different samples is shown in the following graph (Figure 16).

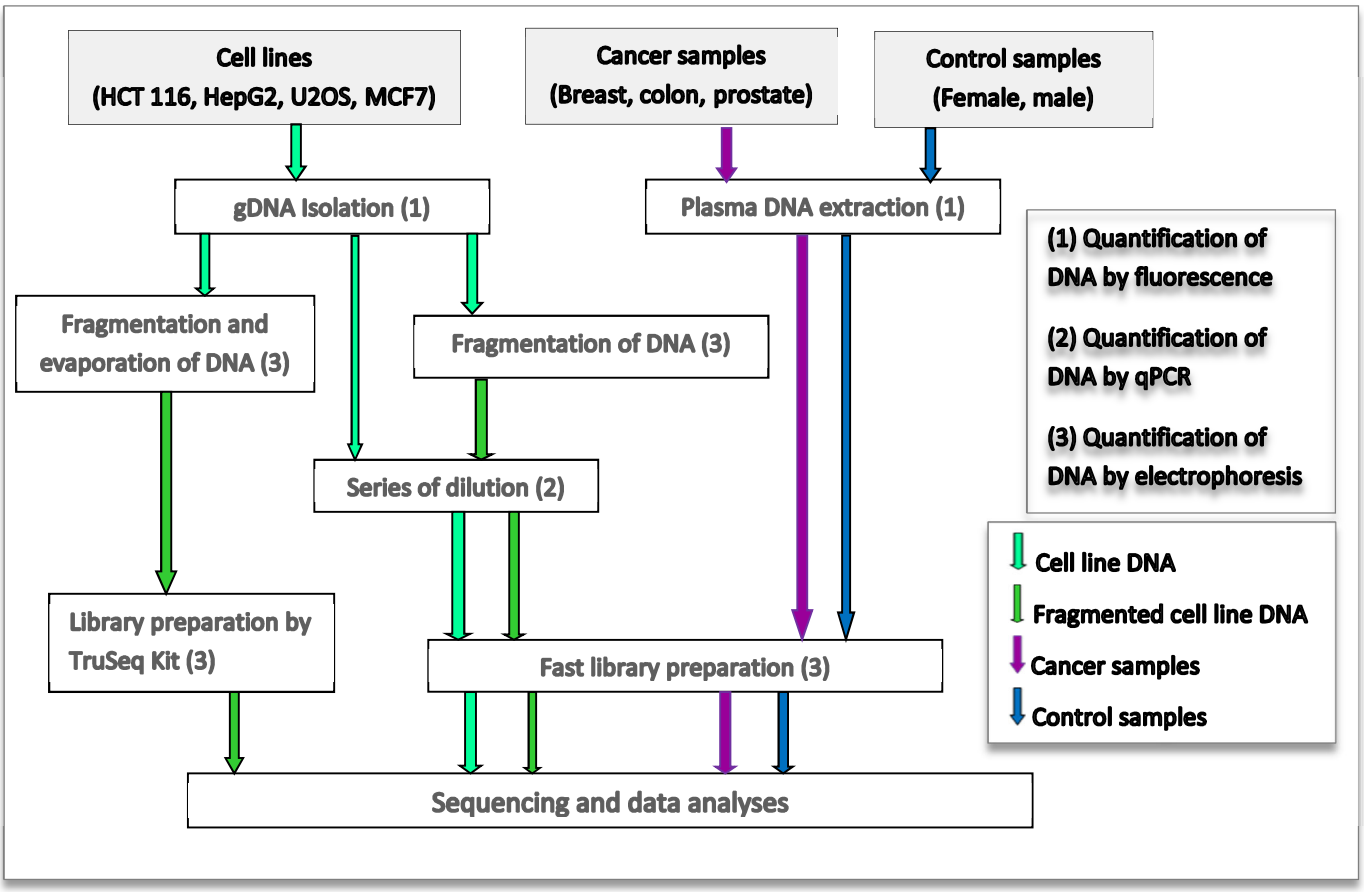


Figure 16: The workflow of the DNA preparation depending on the origin of the DNA.

After the isolation of the gDNA from the cell lines, three different methods were employed. First, the gDNA was fragmented and evaporated and a whole genome library was prepared

with the TruSeq Kit that was later on quantified and sequenced. The second method included an exact quantification by qPCR and an establishing of dilution series of all four cancer cell lines. Those dilution series of the cancer cell lines were amplified with two PCRs. The first PCR used LINE primers that amplified specific regions of the genome and afterwards those regions were amplified with index primers that are necessary for the identification of the samples during sequencing. After those two PCRs the product was again quantified and then sequenced by MiSeq from Illumina. The third employed method for the cell lines was very similar to the second method. Instead of using the long fragments of the DNA, the gDNA was fragmented down to about 250 bp long sections to simulate plasma DNA. After the fragmentation the DNA processed as in the second method explained above.

Once the plasma DNA of the cancer samples was extracted from the blood and quantified, it was directly amplified with the two PCRs as the gDNA from the cell lines. Later the PCR product was quantified, sequenced, and analysed. The samples of the healthy people were processed in the same way as the cancer samples and used as control samples.

The following chapter explains the library preparation, the sequencing and the data analyses.

3.2.1 Cell line HCT 116

The term HCT 116 defines a human male colorectal carcinoma cell line, which is widely used in colon cancer research [Crowley-Weber(2002); Yeung *et al.*(2010)]. This cell line harbours among others an activating *K-RAS* mutation and a *BRAF* mutation, so called oncogenes that are frequently used to predict the resistance of colorectal carcinomas to therapy targeted to the epidermal growth factor receptor [Arcila *et al.*(2011)].

3.2.2 Cell line HepG2

HepG2 cells derived from the liver of a 15 year old male with differentiated hepatocellular carcinoma and have a model chromosome number of 55 [Costantini *et al.*(2013)]. HepG2 cells are often used as hepatocellular carcinoma model and an analysis of genes shows that thousands of genes are significantly down- or up-regulated compared to hepatocytes. As the other cell lines too, they have been grown and harvested at the Institute of Human Genetics.

3.2.3 Cell line U2OS

The cell line U2OS is one of the first generated cell lines [Niforou *et al.*(2008)] and was isolated from a osteosarcoma of the bone of a fifteen year old female [Pontén, Saksela(1967)]. This cell line is chromosomally highly altered, whereas very few normal chromosomes are present. Niforou *et al.*(2008) characterized the U2OS cell line regarding

the genes and their products using proteomics technologies and they found eleven oncogenes related to the cancerous state of this cell line.

3.2.4 CELL LINE MCF7

The abbreviation MCF7 stands for Michigan Cancer Foundation 7, where the cell line was established in the 1970s [Soule *et al.*(1973)]. It names a breast cancer cell line established from a 69 year old female patient, which became a standard model in a variety of laboratories [Levenson, Jordan(1997)]. The cells are derived from the adenocarcinoma of the mammary gland and show a high chromosome number ranging around 82. As one mutation example MCF7 cell lines harbour a *PIK3CA* mutation as well as a high gene copy number alteration [BioMed Central(2005)].

3.2.5 DNA HARVESTING OF CELL LINES

The genomic DNA of all four cell lines was isolated with the QIAamp DNA Micro Kit from QIAGEN according to the user manual.

The cell pellets were diluted with 180 µl of buffer ATL and equilibrated for ten minutes at room temperature. In total 45 µl of proteinase K was added to each cell line pellet, which digests the proteins and releases the nucleic acid. It was left overnight together with the lysate in the tubes and the DNA was harvested the next day using QIAamp Min Elute columns. Therefore, the DNA binds specifically to the silica-membrane of the column and contaminants and inhibitors are completely removed using different washing buffers in the centrifuge (Eppendorf). With the last spin down step the DNA was eluted by the buffer AE (10 nM Tris-Cl and 0.5 mM EDTA pH 9).

After harvesting, the quantity of the isolated DNA was measured by a Qubit® instrument (see 3.2). One part was set aside in the fridge for the following sequencing library preparation; the other part was later on fragmented with a Covaris instrument (see 3.3).

3.2.6 PLASMA DNA FROM BREAST CANCER PATIENTS

The plasma DNA samples of the breast cancer patient were provided from the Institute of Human Genetics and had been collected and processed earlier for previous surveys. A total of 28 breast cancer samples are included in this study, the concentration of the plasma DNA ranged between 16.4 ng/µl to 0.059 ng/µl. A copy number variation profile was established previously using array CGH and Plasma-Seq at the institute, which will be compared later with the results of the used fast sequencing library preparation in the discussion part. Using array CGH seventeen samples have been detected previously as unbalanced and ten samples as balanced (one sample was not analysed by array CGH to date). Otherwise nineteen

samples were classified as unbalanced and four samples as balanced by Plasma-Seq (five samples have not been analysed by Plasma-Seq to date).

3.2.7 PLASMA DNA FROM PROSTATE CANCER PATIENTS

The Institute of Human Genetics also delivered the plasma DNA of the prostate cancer patients. The previous study focused on the development of prostate cancer over time; therefore different groups of patients were available. Patients of group 1 were under active surveillance without surgery, group 2 consists of patients before prostatectomy, group 3 patients are metastatic and patients in the RTX group are treated by radiation therapy. This current study examined four samples of group 1, nine samples of group 2, eleven samples of group 3, and two samples of group RTX. The DNA concentration of the prostate plasma DNA ranged between 106.21ng/ml and 0.74 ng/ml. The Gleason score of the samples is indicated from six to ten.

3.2.8 PLASMA DNA FROM COLON CANCER PATIENT

In addition, a plasma sample of a colon cancer patient was processed with a plasma DNA concentration of 21.51 ng/ml that was obtained for the study of colon cancer markers. The sample was obtained from a highly metastatic patient.

3.2.9 PLASMA DNA FROM HEALTHY PATIENTS

To be able to compare our results to healthy individuals, a control group of 24 female and eighteen male individuals was established. Therefore, those control samples were chosen randomly from the pool of plasma DNA of healthy patient that had been collected for a previous study. The sequencing results of this control group were also used for the calculation of the z-score to determine aneuploidy (see 3.11.3).

3.2.10 PLASMA DNA EXTRACTION

The blood from the patient was collected in the hospital (Landeskrankenhaus Graz) in EDTA tubes and sent directly to the institute. Before the extraction of the plasma DNA the plasma was isolated from the whole blood.

The entire blood was transferred from the EDTA tube into a 15 ml tube and centrifuged (Allegra X-12R Centrifuge) at room temperature for ten minutes at 200 g, afterwards the sample was centrifuged for 1600 g for another ten minutes. Then the supernatant was removed and the rest is filled into another tube that is centrifuged for an additional ten

minutes at 1600 g. The supernatant is then distributed into 1.5 ml tubes to isolate the plasma DNA using QIAcube®.

With the QIAcube up to twelve plasma samples can be processed usually at the same time and fully automated. However, the DNA isolation was carried out with four samples at the same time using per patient in total 3 ml of their plasma to get a higher end concentration of plasma DNA. The purification of the DNA takes about 80 minutes and eliminates the manual processing steps. The machine has a preinstalled protocol for the isolation of plasma DNA that includes the following steps:

(1) Lyse: the plasma sample is lysed in the orbital shaker with the proteinase K.
(2) Incubate
(3) Bind: the lysate is transferred to a spin column in the rotor and the DNA binds to the silica membrane of the column.
(4) Wash: the columns are washed with different buffers (AW1, AW2) to remove contaminants.
(5) Elute: the spin column is transferred to another tube for elution with water.

The plasma DNA extraction is followed by the quantification of the DNA by fluorescent method (see 3.3).

3.3 Quantification of DNA through fluorescence

For a detailed analysis of the dilution series as well as for the amplification through PCR an accurate amount of DNA must be used. Therefore, the quantity of DNA was measured for all genomic DNA samples as well as for the plasma DNA samples of the cancer and the control samples after DNA isolation.

The DNA quantification was performed by Qubit®2.0 flourometer (Life Technologies), because it enables detection of DNA even at low concentrations such as isolated plasma DNA. This technology uses Molecular Probes® dyes which emit signals when bound to specific target molecules and therefore needs to be protected from artificial light during use with a tin foil.

The Qubit® dsDNA BR Assay kit (broad range for 2 - 1000 ng) was employed for the genomic DNA, whereas due to the lower concentration the quantity of the plasma DNA was measured using the Qubit® dsDNA HS Assay kit (high sensitivity with a range of 0,2 - 100 ng). Both Qubit® assays that are utilized with the flourometer use the same general protocol.

The working solution thus mixed the assay reagent to the dilution buffer was mixed with 2 µl of the sample DNA at room temperature in Qubit® tubes. After an incubation time of two minutes when the dye is intercalating in the DNA, the assay was measured from the emitted

signal in the flourometer. In addition two pre-diluted standards were measured from which the Qubit® flourometer calculated the concentration of the samples.

The results of the Qubit® measurement from the genomic DNA were used to check if the isolation was working successfully for the genomic as well as for the plasma DNA. Furthermore, it was used to set up the qPCR for the dilution series.

3.4 Fragmentation of genomic DNA

One part of the genomic DNA was fragmented to approximately 250 bp long sequences to simulate plasma DNA. According to the handbook of Covaris a DNA concentration of 10 ng/µl diluted in 130 µl of buffer is necessary for a proper fragmentation. The isolated DNA was diluted with 1×TE buffer according to the Qubit® quantity results.

The focused-ultrasonicator delivers controlled energy to the sample to shear the DNA into specific fragment lengths while it is computer controlled by SonoLab™ software. The bench-top system operates at high frequencies and is therefore inaudible. The microTUBE focuses the acoustic energy and boosts the inertial cavitation that shear the DNA into fragments with a previous defined length. The diluted DNA samples were fragmented with distributions centred from 250 to 300 bp using the Adaptive Focused Acoustics™ technology. The samples were sheared for 210 seconds in total using the 300 bp method adjusted in the SonoLab™ software.

The samples have been run with the 2100 Bioanalyzer instrument afterwards to check the size distribution of the fragments (see 3.5).

3.5 Quantification of DNA through electrophoresis

After the fragmentation and the amplification by PCR the size distributions need to be characterized to ensure the correct sequence length of the nucleic acids.

The Agilent 2100 Bioanalyzer enabled sizing and quantification of DNA through electrophoresis on a chip. A sample volume of 1 µl was used to analyse the DNA with the Agilent DNA 7500 Kit, which covers a sizing range from 100 to 7500 bp. Up to twelve samples were analysed on one chip with a quantitative range of 0.5 to 50 ng/µl. The chip was loaded with gel in the priming station and the marker, the ladder and 1 µl of the samples were added to the corresponding well. The loaded chip was mixed by vortexing (Bioanalyzer chip vortexer) for one minute and run on the Agilent 2100 Bioanalyzer.

For an accurate run the electropherogram of the ladder was checked considering a flat baseline, the correct identification of the lower and upper marker, and the number of the

twelve high resolved ladder peaks. Afterwards the results of the samples were obtained from the region table.

If PCR products showed the characteristic peaks at the size from 250 up to 350 bp, they were further sequenced.

3.6 Normalisation of dilution series

To achieve a more accurate estimate of concentration to prepare the dilution series of the cell lines the concentration of previous quantity measurements were verified by real time PCR. According to the concentration measurement of the Bioanalyzer for each of the fragmented cell line samples a probe of 50 ng/µl was prepared using nuclease free water for the dilution. The desired concentration was calculated with the following formula:

$$c_1 \times v_1 = c_2 \times v_2$$

c… concentration [ng/µl], v… volume [µl]

With the given concentrations the end-volume v_1 can be calculated.

Based on the results of the Qubit® quantification, DNA samples of 50 ng/µl concentration were prepared that had not been fragmented yet.

3.6.1 Preparation of dilution series

Based on the 50 ng/µl concentration the samples were diluted further in a ratio 1:2 for a total of seven watering steps from 50 ng/µl concentration to 0.4 ng/µl.

For the serial dilution, nuclease free water and some 1.5 ml tubes were necessary. First, 10 µl of nuclease free water was put in all the remaining tubes. Then 10 µl of the 50 ng/µl concentration sample were transferred and mixed into the first tube that contained 10 µl of nuclease free water to cause a 25 ng/µl dilution. From this dilution another 10 µl was transferred to the next tube to produce a 12.5 ng/µl dilution and so on (Figure 17).

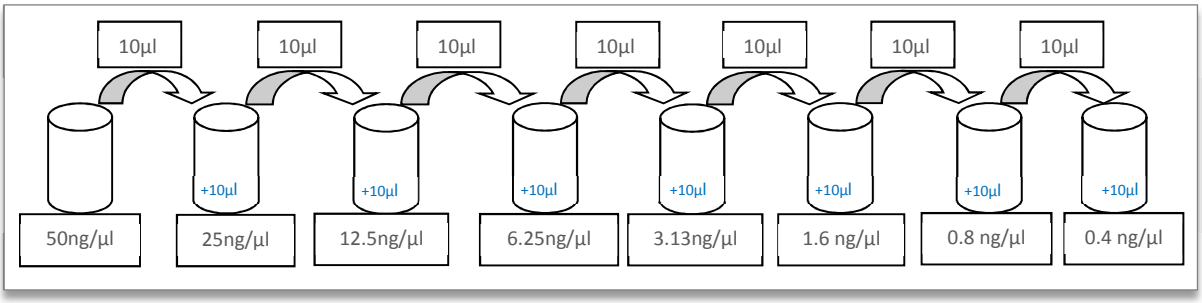


Figure 17: Preparation of dilution series for qPCR.

3.6.2 QUANTITATIVE PCR

The qPCR measured the exact quantity of DNA using a reference human genomic DNA from Promega, Madison. Therefore, a dilution series for the standard was prepared as well. Triplicates were made for all samples to eliminate handling errors in the analyses. The work was performed under the hood and without any artificial light because of the light sensitive dye.

A master mix without the DNA samples was created according to the table 11.

Table 11: Master mix for qPCR

Amount for 1 reaction	Consumable
10 µl	Sybr® Green
1 µl	DNA sample
1 µl	10 µM R1 Primer mix
8 µl	nf. water

After 19 µl of the master mix was loaded into every well of the reaction plate, 1 µl of the DNA of the dilution series was added and the plate was sealed with the adhesive film. The plate was then loaded in the real time PCR instrument that is computer controlled by the StepOne™ software. The concentrations of the reference human genome were set as the standard and the concentration of the cell lines were calculated based on the standard. The qPCR followed the temperature program in table 12, later the data was analysed with the StepOne™ software.

Table 12: Temperature program for qPCR

FAST 1		
temperature	time	
95°C	5 min	
95°C	30 s	A total of 25 cycles
60°C	30 s	
10°C	∞	

3.6.3 STATISTICAL ANALYSIS OF QUANTITATIVE PCR DATA

Once the qPCR was finished the melt curve and the standard curve were checked using the StepOne™ software. In the melt curve all samples should have similar temperature behaviour and the C_T value of all samples should lie on the standard curve for a reliable and accurate measurement of the DNA quantity. Figure 18 shows the melting curve and the

standard curve of the HepG2 dilution series. The blue unknown sample indicates the cell line and the red standard the Promega dilution samples.

The amplification plot was examined and outliers were omitted in the plate layout (Figure 19). The filtered data was exported in an excel file and the ratio of Promega to cell line DNA concentration was calculated. Therefore, the mean of the quantity of the filtered sample with the same dilution concentration was calculated. For every dilution, the ratio of the Promega DNA's quantity to quantity of the cell line DNA was calculated. The mean of all ratios was used to dilute the 50 ng/µl cell line samples to the corrected concentration and the dilution series were prepared (see 3.7). Table 13 shows the calculation of the Promega HepG2 ratio as example how to determine the ratio.

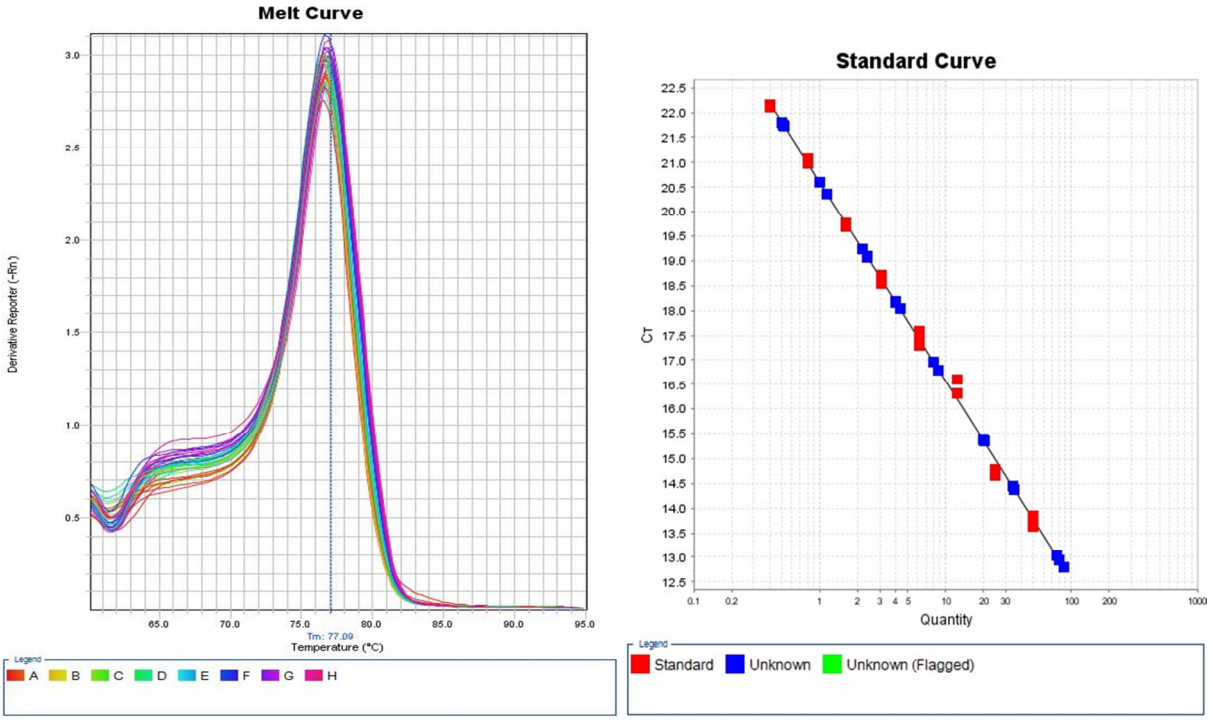


Figure 18: Melt curve and standard curve of qPCR.

Table 13: Calculation of the Promega- HepG2 ratio for the serial dilution

Promega	HepG2	Ratio
50	81.378	0.614
25	34.522	0.724
12.5	20.048	0.624
6.25	8.433	0.741
3.13	4.131	0.756
1.56	2.315	0.675
0.78	1.224	0.638
0.39	0.513	0.762
Mean		**0.692**

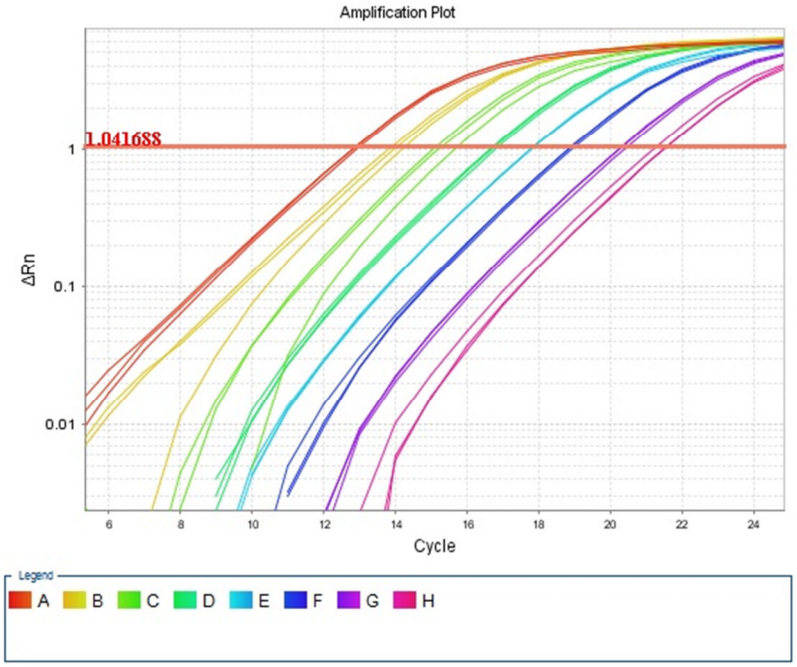


Figure 19: Amplification plot of HepG2 dilution series.

3.7 Preparation of dilution series of cell lines

To detect the sensitivity of the fast sequencing library preparation method eight series of dilutions were created. The cell lines HCT116 and HepG2 were mixed with the male human genomic DNA provided by Promega (Madison), whereas the cell line U2OS and MCF7 were mixed with the female human genomic DNA from Promega (Madison). Both genomic standards came from multiple anonymous donors. Later on, all four fragmented cell lines were mixed with the female human genomic DNA, as the X- and Y-chromosome were not included in the statistical analysis of the sequenced genome of male cell lines.

For example, HCT116 100% consisted of only cell line DNA, whereas HCT116 50% consisted half of Promega DNA and half of cell line DNA and HCT116 0% has no cell line DNA only Promega DNA and should therefor show no aneuploidy in the results (Table 14).

Table 14: Dilution series of cell lines

	Cell line	Standard DNA
100%	10 µl	0 µl
75%	7.5 µl	2.5 µl
50%	5 µl	5 µl
25%	2.5 µl	7.5 µl
20%	2 µl	8 µl
15%	3 µl	17 µl
10%	2 µl	18 µl
5%	1 µl	19 µl

2.5%	1.25 µl	48.75 µl
1%	0.5 µl	49.5 µl
0%	0 µl	10 µl

3.8 FAST PCR

The plasma samples as well as the dilution series were amplified with two PCRs to be later on sequenced and analysed. A maximum of 50 ng of DNA was utilized for the first PCR if achievable, for the plasma DNA amplifications a lower amount was enough to get a product. The first PCR uses LINE specific primers that amplify randomly distributed regions throughout the genome. The PCR product is then cleaned with magnetic beads that bind the DNA and determinants can be washed out with ethanol. After elution of the DNA a second PCR was set up and the sequences of the first PCR were used as a template on which different index sequences were attached during the second PCR to identify the pooled samples during sequencing.

3.8.1 PCR 1

The consumables for one gDNA reaction are listed in the table 15. The sequences of the bases of the forward and reverse primer are attached in the appendices.

Table 15: Master mix for PCR 1

Consumable	Amount
DNA (50 ng)	1 µl
Nuclease free water (Promega, Madison)	34.5 µl
5× Phusion HF Buffer (ThermoScientific)	10 µl
10 µM forward and reverse Primer (R1-for a and R1-Rev) (Microsynth)	2.5 µl
10 mM dNTPs (Invitrogen)	1 µl
2U Phusion Hot Start II Polymerase (ThermoScientific)	1 µl
Total	50 µl

Due to the lower concentration of DNA in plasma 5 µl of DNA instead of 1 µl for all the cancer samples were amplified and the amount of water was reduced to 30.5 µl per

reaction. The control plasma DNA showed even a lower concentration, therefore 15 µl of plasma DNA were used in the first PCR.

A master mix was prepared according to the table 16 seen above and the PCR tubes were loaded with the master mix. The DNA was added to each tube and mixed by pipetting. After a brief centrifuge the thermo cycler was run with the following program (lid was heated to 100°C) as shown in the table 16.

Table 16: Temperature program for PCR 1

FAST 1		
temperature	time	
98°C	2 min	
98°C	10 s	A total of 5 cycles
57°C	2 min	
72°C	2 min	
4°C	∞	

3.8.2 PCR CLEAN-UP

After the first amplification was accomplished the PCR product needed to be purified with magnetic beads. The consumables for one reaction for the clean-up are listed in the table 17.

Table 17: Consumables for the clean-up

Consumable	Amount
70 % Ethanol	400 µl
AMPure XP beads (Beckman Coulter)	70 µl
1× TE Buffer	12 µl

50 µl of each PCR product was transferred to a 1.5 ml labelled tube and mixed with 70 µl of AMPure beads and vortexed thoroughly for five seconds. After incubation for ten minutes at room temperature the DNA was bound to the magnetic beads and the tubes were placed on the magnet for another five minutes. Then the supernatant was removed and the bead pellet was washed with fresh prepared 70% ethanol two times. After the second wash the pellet was dried at room temperature for ten minutes on the magnet. Once dry the pellet was resuspended in 12 µl of 1× TE buffer by pipetting and incubated again for two minutes to remove the DNA from the beads. Finally, the tubes were placed again on the magnet and incubated for two minutes. With the tubes still on the magnet, 10 µl of supernatants were removed carefully to a set of fresh, labelled tubes and set aside for the second PCR.

Figure 20 shows the process overview of the magnetic beads from the binding of the DNA, the wash steps to the elution.

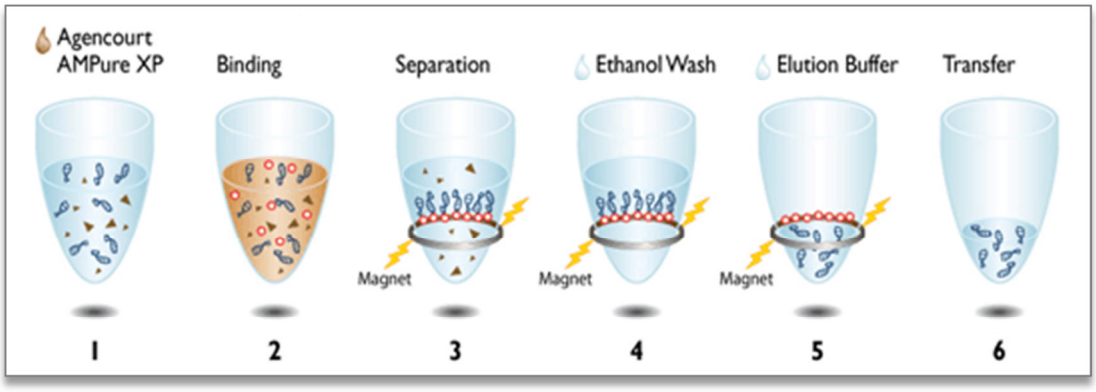


Figure 20: Process overview of magnetic beads: 1. Add beads to PCR reaction; 2. Magnetic beads bind to DNA; 3. Separation of DNA bound to beads from contaminants; 4. Washing DNA with ethanol; 5. Elution of DNA sequences from beads; 6. Transfer DNA away from beads into a new tube [Beckman Coulter (2009)].

3.8.3 PCR 2

The elution of the first PCR was used directly for the second round of amplification. The materials for the PCR 2 were identical with that one from the PCR 1 and the consumables for the second PCR are shown in the table 18. The sequences of the forward and reverse index primers are attached in the appendices. The primers annealed to one site that was introduced by the first round primers and contained the 5' sequences that is necessary for hybridization to the sequencing flow cell later on.

Table 18: Master mix for PCR 2

Consumable	Amount
Elution of first PCR product	10 µl
Nuclease free water (Promega, Madison)	25.5 µl
5× Phusion HF Buffer (ThermoScientific)	10 µl
10 µM forward Primer (R2-For, MicroSynth)	1.25 µl
10 µM reverse Primer (R2-RevA, MicroSynth)	1.25 µl
10 mM dNTPs (Invitrogen)	1 µl
2U Phusion Hot Start II Polymerase (ThermoScientific)	1 µl
Total	50 µl

All reagents were mixed together as listed in the table 18 in a PCR tube. The DNA was added to each tube and mixed by pipetting. After a brief centrifuge the thermo cycler was run with the following program (lid was heated to 100°C) as shown in the table 19.

Table 19: Temperature program for PCR 2

FAST 2		
Temperature	time	
98°C	2 min	
98°C	10 s	A total of 15 cycles
65°C	15 s	
72°C	15 s	
4°C	∞	

To augment the DNA concentration of the plasma DNA the cycles were increased from fifteen to eighteen labs for all plasma samples.

The second PCR was followed again by a purification using the magnetic beads as in 3.8.2. After elution of the PCR amplicons from the AMPure XP beads they were quantified by electrophoresis using the Agilent 2100 Bioanalyzer (see 3.5) and then loaded on the sequencing instrument (see 3.10).

3.9 WHOLE- GENOME LIBRARY PREPARATION

The TruSeq Kit from Illumina was utilized to establish four whole-genome libraries of the four cell lines HepG2, HCT116, MCF7, and U2OS.

First, the DNA needed to be fragmented into short sequences and evaporated to a specific volume. The library preparation involved several sequential steps: after the shearing and evaporation an end repair of the sequences was performed followed by 3'end adenylation that added an A-tail to the DNA sequence to avoid blunt ends. In the next step adaptors were ligated to the A- tail of the fragments that were finally enriched by multiple PCR cycles (Figure 21). The PCR products are then quantified and loaded on the sequencing instrument

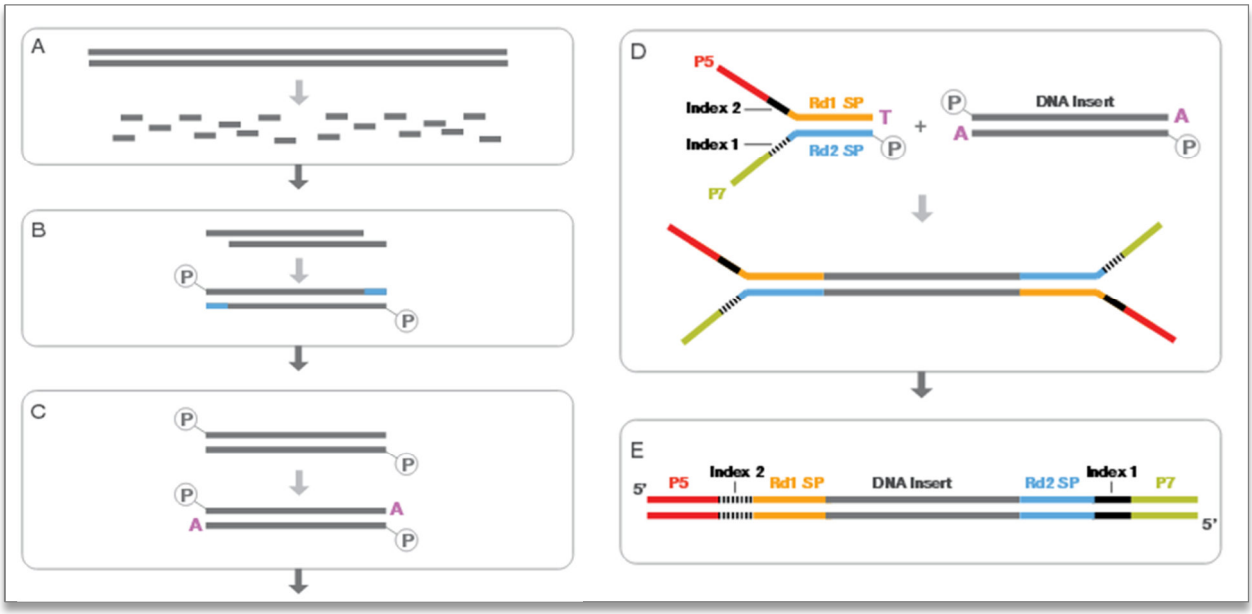


Figure 21: Library preparation; A. genomic DNA is fragmented; B. Blunt ends are generated; C. A- base is added; D. Index adapters are ligated to both ends; E. Product is amplified and then sequenced [Illumina].

3.9.1 FRAGMENTATION AND EVAPORATION

The genomic DNA samples were fragmented with the Covaris machine as the previous plasma samples (see 3.3). Therefore, the same consumables and materials were used. The instrument was run for 80 s per sample and set to the 300 bp method.

After the fragmentation samples were evaporated to a volume of 50 µl in the Concentrator Plus (Eppendorf) instrument. The concentrator function of this instrument enabled a fast concentration of samples using both, vacuum and centrifugation in the rotor. Sheared DNA

samples were transferred in normal 1.5 ml tubes and were placed in the rotor chamber after the mandatory warming up phase. The instrument was turned on for around 35 to 40 min while the method was set to V-AQ (suitable for aqueous solution).

3.9.2 Performing end repair

Sheared ends were repaired using the TruSeq DNA Sample Prep Kit. The samples were processed exactly according to the user manual of Illumina. After the incubation of the sheared fragments with the end repair mix for 30 minutes on the thermal cycler at 30°C, the beads were added to bind the nucleic acids. The samples were incubated with the beads for fifteen minutes at RT to bind the fragments to the beads. Then the samples were placed on the magnet for twenty minutes and afterwards the supernatant was removed carefully without disturbing the beads. After two wash step with 80% of ethanol that removed the rest of the determinants, the DNA was dried at RT for twenty minutes. The dried pellet was resuspend with 17.5 µl of resuspension buffer. Therefore, the DNA binding with the beads was interrupted and the DNA fragments were eluted.

3.9.3 Adenylation of 3'ends

For the adenylation of the 3'end of the fragments an A- tailing mix was used according to the user manual that was provided by Illumina.

The samples were incubated with 12.5 µl of A-tailing mix for 30 minutes at the thermal cycler at 30°C. It was necessary to proceed immediately after the incubation to the next step that included the attachment of adapters.

3.9.4 Ligation of adapters

The ligation mix was added together with the DNA adaptor indexes to the adenylated product for ten minutes. After the adaptors were ligated, the process was inhibited by the stop ligation buffer. With two washing cycles using magnetic beads and 80% ethanol all contaminants were removed again. After air-drying the pellet for twenty minutes on the magnet it was resuspended in 22.5 µl of resuspension buffer provided by Illumina.

3.9.5 Enrichment of DNA fragments

The thermal cycler was programmed according to the instructions of Illumina for fifteen cycles (Table 20). The prepared DNA fragments were mixed with the PCR primer cocktail as well as with the PCR master mix and passed in the thermal cycler to amplify the DNA fragments of the whole genome.

Table 20: Temperature program for DNA enrichment

TruSeq DNA enrichment		
Temperature	time	
98°C	30 s	
98°C	10 s	A total of 15 cycles
60°C	30 s	
72°C	30 s	
72°C	5 min	
10°C	∞	

After the PCR, the product was cleaned with the magnetic beads and 80% ethanol on the magnet stand. In the end the dried pellet was resuspended with 32.5 µl of Resuspension buffer from Illumina and quantified by electrophoresis using 2100 Bioanalyzer using the Agilent DNA 7500 Kit.

3.10 SEQUENCING OF SAMPLES

Parallel sequencing of up to 24 samples was realized with the MiSeq technology by Illumina. This bench top instrument combines sequencing and data analysis in one instrument. Before sequencing the samples were denatured and diluted.

3.10.1 PREPARING LIBRARIES FOR SEQUENCING

Samples with different indices were diluted to 4nM libraries and pooled together. The concentration of 4 nmol/µl was calculated based on the Bioanalyzer results.

The MiSeq reagent Kit contains the hybridization buffer, the reagent cartridge and a bottle of PR2 solution. Meanwhile the hybridization buffer HT1 that was used to dilute the denatured libraries was set aside to thaw. The reagent cartridge was thaw using a room temperature water bath for approximately 40 minutes. The NaOH was prepared freshly for denaturing the 4nM libraries for cluster generation in the MiSeq instrument. Therefore 980 µl of water was mixed with 20 µl of 0.1 N NaOH under the hood.

The pooled 4 nM DNA libraries were denatured into single strands with the NaOH for five minutes at room temperature. The HT1 buffer was added to stop the reactions (Table 21) and to dilute the denatured libraries to a final concentration of twelve pM (Table 22). The

denatured and diluted DNA was placed on ice until the samples were loaded onto the MiSeq reagent cartridge.

Table 21: 4nM library denaturation

Library denaturation		
Consumables	Amount	
Pooled DNA	5 µl	Incubate for 5 min at RT
0,2 N NaOH	5 µl	
HT 1	990µl	Dilute to 20 pM

Table 22: Dilute denatured DNA to 12 pM

Library dilution	
Consumables	Amount
20 pM denatured DNA	360 µl
HT1	240 µl

All samples that have been generated with the same fast method consisting of the two PCRs were pooled. The fast method generated low diversity libraries that means that a significant number of the reads have the same number due to the LINE primers used in the first PCR. This causes a shift in the base composition as the reads are not random any more. A higher concentration of a whole genome library prevented this shift and helped to balance the lack of sequence diversity.

The whole genome libraries of the four samples from the cell line were used to increase the diversity of the pooled library. They were therefore denatured and diluted separately according to the table 21 and 22. After the dilution the sample library was combined with the whole genome library (WGA) as shown in table 23.

Table 23: Combine sample library and whole genome library

Combination of libraries	
Consumables	Amount
WGA library	360 µl
Fast library pool	240 µl

3.10.2 SPIKE-IN OF PHIX CONTROL

Illumina recommends increasing the diversity of the libraries by sequence them not only with other diverse libraries like WGA libraries but also with the PhiX control. Therefore, some samples were spiked in with a PhiX control, which is commonly used for Illumina sequencing if a low diversity library occurs. The PhiX control is derived from the well characterized bacterial genome of the same name.

2 µl of the 10 nM PhiX library was diluted with 3 µl of nuclease free water and denatured with a freshly prepared 0.2 N NaOH. After denaturation with NaOH and dilution with the hybridization buffer HT1 to a 12 pM library, the PhiX control was combined with the pooled samples (Table 23).

3.10.3 PERFORMING A SEQUENCING RUN

After the reagent cartridge was prepared for use, the pool of indexed libraries was loaded and a volume of 600 µl of the sample libraries was pipetted into the reservoir.

A sample sheet for the MiSeq was written using the Illumina Experiment Manager including the names of the pooled samples and their indices. The sample sheet stored information needed to set up, perform, and analyse the sequencing data. From the MiSeq Control software (MCS) interface, “Sequence” was selected to start the run set up steps. The flow cell was rinsed with nuclease free water and dried with a tissue. The flow cell, the PR2 bottle and the reagent cartridge with the pooled samples were loaded on the MiSeq instrument and the waste bottle was emptied.

From the software interface, “Start Run” was selected to start the sequencing run after reviewing the run parameters. A 150 cycle run was chosen for all samples, because it was faster than the 300 cycle run and had enough accuracy for the experiment. At the end of the run, two times 150 cycles were analysed.

Once the run started, a sequencing screen opened which provided a visual representation of the run including cluster density, cluster passing filter and quality scores. A high quality score for Q30 was desired which means that the probability of a wrong base is one in 1000. A cluster passing filter of 80 % was normally achieved. Therefore 80 % of all clusters were able to pass the chastity filter, which measured the quality.

3.11 Analysis of sequencing data

After the MiSeq run, the data was analysed according quality standards and then further processed to identify aneuploidy within the samples. Parallel sequencing generates a massive amount of data that needs to be aligned to a standard human genome. Before the data was used to calculate any statistical data it was filtered and normalised.

3.11.1 Sequence alignment

Using linux tools the sequencing reads were counted per chromosome for the dilution series. For the cancer samples as well as for the dilution series a specific algorithm was used to count the reads per chromosome arm.

All sequence tags that passed the Illumina chastity filter were saved in a FASTQ file. The files had been separated by sample. Therefore, each file contained the sequences for a particular sample and read. The reads were aligned to the human genome (hg 19) using the BWA algorithm (Burrows-Wheeler Alignment) and saved in an aligned file [Li, Durbin (2009)]. The BWA tool, which is based on the Burrows-Wheeler Transform, aligns efficiently short sequencing reads against a long reference sequence, allowing mismatches and gaps. The read alignment package supports the Illumina sequencing machines as well as AB SOLiD instruments.

```
bwa aln –f “AlnFile” -t 4 ~/RefSeq/hg19_070510/hg19 “FastQ_File”
```

The aligned file was converted to a human readable SAM-file (Sequence Alignment/Map) and downstream analyses can be achieved with the open source SAM tools. SAM format allows simply and flexible storing of large nucleotide sequence alignments in a compact file size [The SAM/BAM Format Specification Working Group (2013)]. In the format each alignment line represents the linear alignment of a segment with mandatory fields that hold information such as the name, the reference sequence name, and the mapping quality. The fifth column of the fields contains the mapping quality, which equals - 10 $\log_{10}$ of probability that mapping position is wrong and is rounded to the nearest integer. Low quality alignments, which had a mapping quality lower than fifteen, were filtered out using common linux tools (cat, awk).

```
cat “SamFile” | awk '$5 > 15' > “NewSamFile”
```

After quality masking, the remaining alignments per chromosome were counted using wordcount (wc) and piping operator |.

```
Cat “SamFile” | awk ´$3 == “chr1” ´ | wc –l > CountSamFile
```

The counted alignments per chromosome were saved in Excel and normalised.

3.11.2 NORMALISATION

Every sample generated a different amount of total reads due to stochastic and experimental variations. Therefore, it was necessary to normalize the data to make a significant comparison between the samples.

Table of read counts were generated for each sample and a read-count ratio was calculated (reads per chromosome/ total reads) for every chromosome.

3.11.3 IDENTIFICATION OF ANEUPLOIDY STATUS

For identification of the aneuploidy status the z-score statistic was applied on the normalised data, being a common method of determining the aneuploidy status.

First the mean and the standard deviation of the read-count ratios of the group of reference samples were determined. The female control samples were the reference group for all female cell lines (U2OS and MCF7) and breast cancer patients, whereas the male control samples formed the reference group for all male cell lines (HCT116 and HepG2) and the colon and prostate cancer patients.

The z-score was calculated by the subtraction of the mean read-count ratio of control samples from the read-count ratio of the sample x divided through the standard deviation read-count ratio of control samples.

$$z_x = \frac{x - \mu_x}{\sigma_x}$$

σ_x… standard deviation read-count ratio of control samples, μ_x… mean of the read-count ratio of the control samples

As all samples were standardised in this way, two thresholds were set to detect outliers and they were easily detected:

- $z_x < -3$ losses of chromosomes or parts of the chromosomes
- $z_x > +3$ gains of chromosomes or parts of the chromosomes

3.11.4 VISUALISATION OF Z-SCORES

In addition to the tables of z-scores, a graphical interpretation of the z-scores, so called heatmaps, was implemented in R for every dilution series and the plasma DNA samples. The

excel file with the calculated z-scores were converted to a .txt file. Using R version 3.0.2 following code was used to create the heatmaps without reordering of rows and columns:

```
require(graphics); require(grDevices)

genes<- read.table("C:/Users/Tesi/Desktop/Masterthesis/
R_Files/breastcancershadesarm.txt", header=TRUE)

row.names(genes) <- genes$Name

genes <- genes[,2:49]

genes_matrix <- data.matrix(genes)


colors = c(seq(-110,-50,length=4),seq(-50,-
3,length=20),seq(-3,3,length=2),seq(3,50,length=20),
seq(50,110,length=4))


genes_heatmap <- heatmap(t(genes_matrix), Rowv = NA, Colv =
NA, col = colorRampPalette(c("darkblue","blue","gray93",
"red","red3"))(49), breaks = colors, scale="none", margins =
c(7,5), xlab = "Patient samples", ylab =  "Chromosome", main
= "zScores of breast cancer samples ")

legend("topright", fill= c("darkblue","blue","gray97","red",
"red4"), legend= c("< -50","-50 to -3","-3 to 3", "3 to 50",
"> 50"), bty != "n", bg =c("lavenderblush"), text.col=
c("gray15"), text.font= 2,title=c("Z-score"), title.col =
c("gray25"))
```

4. RESULTS

4.1 QUANTITY AND QUALITY CONTROL

All samples were quantified using electrophoresis and the size distribution was examined for the fragmented samples before proceeding to the library preparation step.

The fragmentation of the cell lines resulted in an average size of 250 to 300 bp and an even size distribution. Figure 22 shows the size distribution with an average size of 258 bp of the fragmented cell line U2OS. The concentration was used to set up the real time PCR to determine an accurate dilution series, with which a sequencing library was then prepared using the adapted FAST-SeqS technique.

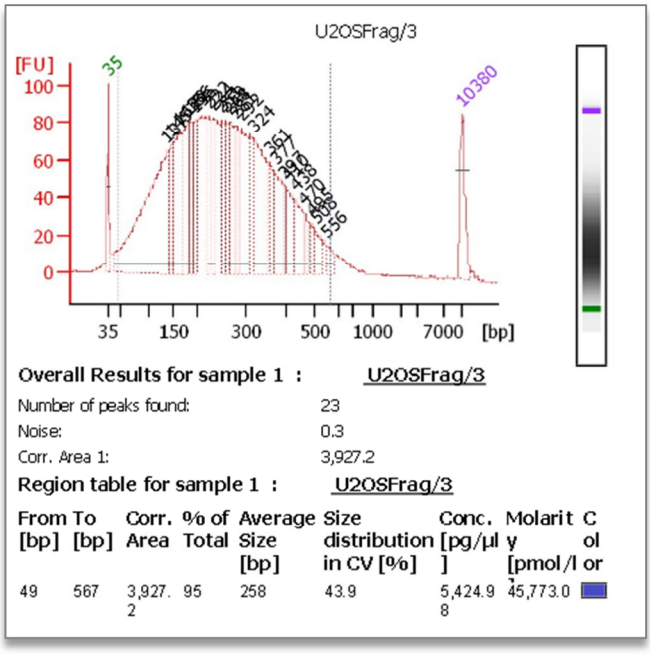


Figure 22: Size distribution of fragmented cell line U2OS.

The prepared libraries of control, cancer and cell line samples were examined according to the size distribution and quantity. Figure 23 shows the characteristic peaks for the breast cancer sample B41 and the fragmented cancer cell line sample HepG2 100%. The FAST-SeqS library preparation results in a characteristic peak, which has the maximum at the size of 250 bp and covers a region from about 180 bp up to 350 bp. The molarity of each sample that was obtained with the electrophoresis was further used to dilute them to a 4 nM library.

The diluted libraries were sequenced in the MiSeq instrument in which every run consisted of 150 cycles in this study. Phred scores defined the accuracy of DNA sequencing for every

cycle indicating the probability that a given base is not called correctly. If the Phred quality score equals 30 (Q30), the run has a base call accuracy of 99.9 % with an incorrect base call one in 1000 times. The vast majority of basis of all samples, on which data filtration and analysis were conducted later on, reached a higher sequencing quality than Q30. Therefore, almost all reads were considered to be perfect without errors and ambiguities.

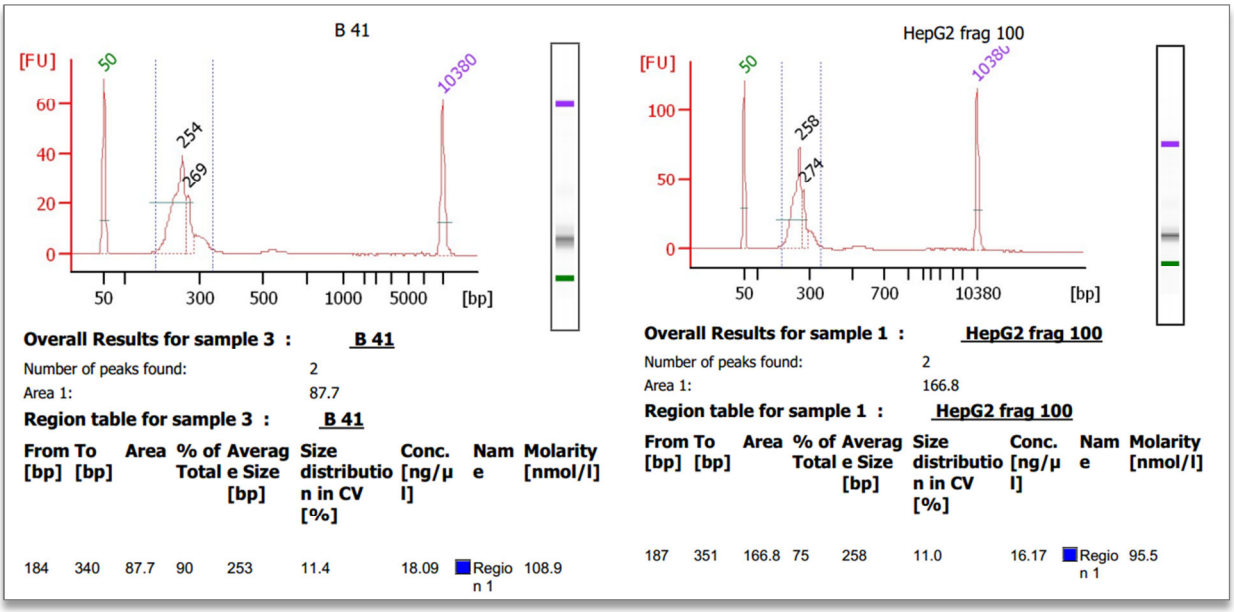


Figure 23: Size distribution of a breast cancer sample and a sample of the dilution series of HepG2 after fast library preparation.

For each sample, reads obtained from the sequencing run with quality score of more than 30 were mapped to the human reference genome hg 19. And to determine if the percentage of mapped reads to the chromosomes derived from plasma of cancer patients differs from the mapped reads derived from plasma of non-malignant individuals, z-scores were calculated.

4.2 Control samples

A common method for detection of chromosomal copy number changes is by comparison of z-scores. This analysis is based on the mean and standard deviation of read distribution within a chromosome of control samples. For the study, plasma DNA samples of 24 women and eighteen men without any indication of a malignant disease were used. So first, defined templates of control samples were amplified and tagged with a unique identifier DNA sequence. The prepared amplicons were sequenced and z-scores of each chromosome and each chromosome arm of every sample were calculated. The total read counts between the controls, which were obtained from the sequencing run, showed a very low variation (see appendix). Furthermore, only few chromosomes arms, short (p) arms and long (q) arms, showed a z-score above 3 or below -3 for the female controls (Figure 24) as well as for the

male controls (Figure 25). The low variation and z-scores of the whole chromosomes as well as for the chromosome arms proved the sequencing method as robust and reproducible.

A mean of 234,949 reads per sample for female and a mean of 186,528 reads per sample for male controls were achieved from the sequencing run. Standard deviation and coefficient of variation were calculated first for each chromosome (Table 24) and then for each chromosome arm (appendix). In addition, the coefficients of variation, which is defined as the ratio of the standard deviation to the mean, were calculated to show the extent of variability in relation to the mean of the control samples. The coefficient of variation was relatively low for the chromosomes and chromosome arms indicating robustness of the sequencing results.

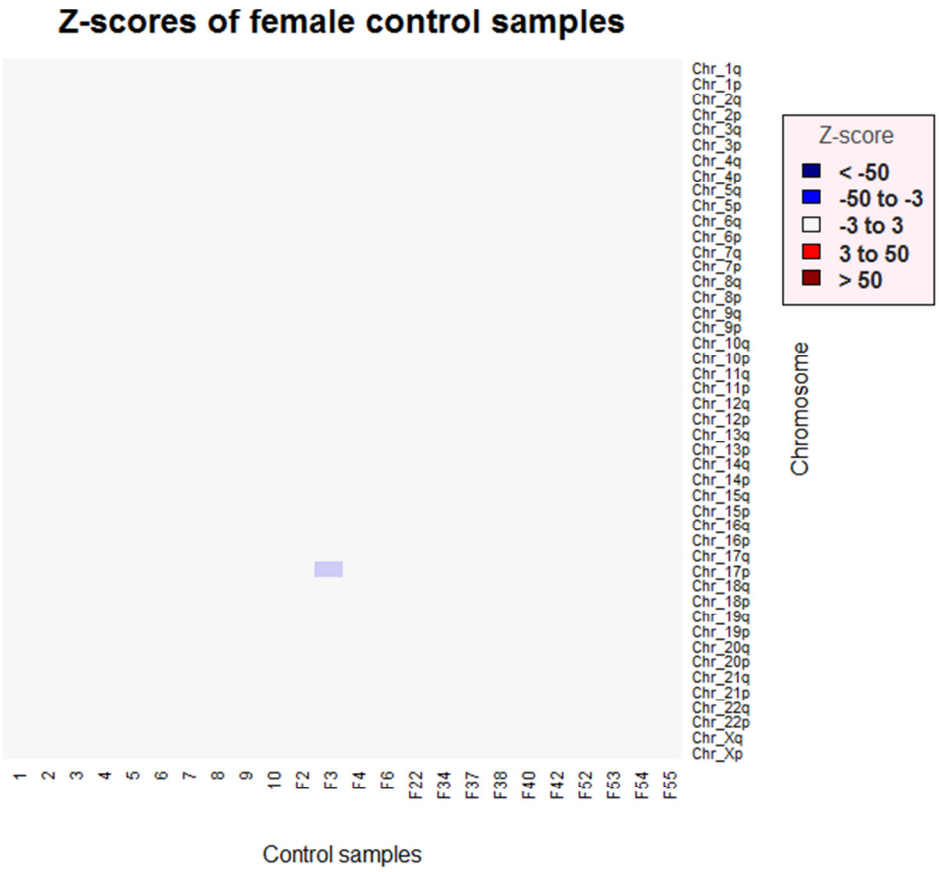


Figure 24: Copy number analysis of female control samples for each chromosome arm. Z-scores above +3 are marked with red colour indicating gains and z-scores below -3 indicate losses marked with blue colour.

Figure 25: Copy number analysis of female control samples for each chromosome arm. Z-scores above +3 are marked with red colour indicating gains and z-scores below -3 indicate losses marked with blue colour.

Table 24: Mean and standard deviation of female and male control samples.

Chromosome	male			female		
	Mean	SD	CV	Mean	SD	CV
Chr_1	0.0707	0.0008	0.0114	0.0677	0.0014	0.0210
Chr_2	0.0870	0.0004	0.0046	0.0830	0.0008	0.0093
Chr_3	0.0769	0.0007	0.0086	0.0750	0.0013	0.0169
Chr_4	0.0835	0.0013	0.0155	0.0785	0.0011	0.0141
Chr_5	0.0754	0.0009	0.0120	0.0704	0.0007	0.0095
Chr_6	0.0673	0.0005	0.0076	0.0646	0.0009	0.0134
Chr_7	0.0558	0.0004	0.0076	0.0537	0.0008	0.0152
Chr_8	0.0568	0.0005	0.0088	0.0540	0.0006	0.0105
Chr_9	0.0382	0.0004	0.0111	0.0369	0.0003	0.0094
Chr_10	0.0448	0.0005	0.0115	0.0432	0.0003	0.0059
Chr_11	0.0462	0.0011	0.0244	0.0429	0.0010	0.0231
Chr_12	0.0505	0.0009	0.0177	0.0487	0.0009	0.0183
Chr_13	0.0383	0.0006	0.0150	0.0371	0.0007	0.0197
Chr_14	0.0287	0.0003	0.0097	0.0271	0.0008	0.0305
Chr_15	0.0219	0.0005	0.0225	0.0215	0.0005	0.0238
Chr_16	0.0221	0.0004	0.0177	0.0208	0.0007	0.0359
Chr_17	0.0184	0.0009	0.0499	0.0190	0.0008	0.0430

Chr_18	0.0259	0.0005	0.0184	0.0243	0.0007	0.0277
Chr_19	0.0102	0.0007	0.0655	0.0088	0.0009	0.1040
Chr_20	0.0149	0.0004	0.0281	0.0144	0.0005	0.0330
Chr_21	0.0161	0.0002	0.0137	0.0154	0.0004	0.0262
Chr_22	0.0053	0.0003	0.0510	0.0053	0.0003	0.0582
Chr_X	0.0452	0.0008	0.0174	0.0875	0.0016	0.0184

The genome wide z-score was calculated using the z-scores of the chromosome arms and are shown in the table 25. First the sums of square of all chromosome arms were calculated for every female and male control sample. The sum of square was normalised using the mean and standard deviation of the other control samples of the same gender. Only one female control showed a genome-wide z-score higher than three, which resulted from loss on chromosome arm 17p. As expected, all other samples showed a genome-wide z-score lower then tree.

Table 25: Genome-wide z-scores of female and male control samples. Genome-wide z-scores above three are highlighted in red colour.

Female controls	Sum of Squares	Genome-wide z-scores	Male controls	Sum of Squares	Genome-wide z-scores
1	56.43	0.53	**A**	126.32	2.65
2	68.74	0.94	**F**	118.50	2.27
3	52.84	0.42	**H**	16.60	-1.23
4	16.79	-0.77	**I**	37.66	-0.53
5	16.53	-0.78	**J**	25.09	-0.94
6	20.16	-0.66	**M4**	43.32	-0.35
7	20.61	-0.65	**M5**	32.08	-0.71
8	25.35	-0.49	**M6**	77.79	0.73
9	26.74	-0.45	**M10**	37.38	-0.54
10	19.74	-0.68	**M11**	47.76	-0.21
F2	116.93	2.53	**M19**	63.13	0.27
F3	136.46	3.18	**M27**	47.44	-0.22
F4	45.43	0.17	**M28**	23.83	-0.98
F6	51.92	0.39	**M33**	45.75	-0.28
F22	23.22	-0.56	**M34**	23.31	-1.00
F34	22.85	-0.57	**M43**	76.79	0.70
F37	18.43	0.74	**M44**	37.57	-0.53
F38	25.07	-0.50	**M45**	102.49	1.99
F40	35.64	-0.15			
F42	21.20	-0.63			
F52	41.82	0.05			
F53	37.77	-0.08			
F54	28.97	-0.37			
F55	36.35	-0.13			

4.3 CANCER SAMPLES

For this study, plasma DNA samples of 28 women diagnosed with breast cancer, 27 patients with prostate cancer and one patient suffering from colon cancer were analysed by the adapted FAST-SeqS method. Z-scores of each chromosome arm of every sample, short (p) arms and long (q) arms, were calculated using the previously determined mean and standard deviation of read distribution of the female and male control samples.

The short (p) arms of some acrocentric chromosomes including chromosome 13, 14, 15 and 22 were excluded for z-score determination of all cancer samples. An average read number of 340,010 for the breast cancer samples and 173,593 for the prostate cancer samples were obtained. The colon cancer sample had a read number of 160,077. The proportion of reads presenting each chromosome arm was calculated for each sample by dividing the total read counts that are mapped to that arm by the total number of reads from all chromosome arms. This normalised ratio then was used to calculate the z-scores of every chromosome arm and of the entire genome using the mean and SD of the control samples of the same gender.

4.3.1 BREAST CANCER

The z-scores of the chromosome arms of breast cancer patients that were previously diagnosed by array CGH as balanced are provided in the figure 26. Array CGH is a method for analysing copy number variations by comparing two DNA samples of different origin such as tumour genome with normal reference genome. After cohybridization of the two different fluorescent labelled DNA samples to a DNA microarray the relative fluorescence intensities of each hybridized fluorophore is quantified and a ratio of the intensities is calculated, which is proportional to the ratio of copy numbers of DNA sequences in the test and reference genomes. If the log2 ratios of the fluorescent probes is near to zero, the tested genome is considered to be balanced; if there is an altered ratio this indicates a loss or gain of the tumour DNA. Using the principles of competitive fluorescence in situ hybridization unbalanced chromosomal abnormalities can be detected genome wide with high resolution as low as 100 kilobases. However, balanced chromosomal abnormalities for instance inversions or reciprocal translocations can not be detected, as they do not affect copy number.

In figure 26, z-scores above three obtained with the Fast sequencing method signify a gain of the chromosome arm and are emphasised with different shades of red. The darker the colour the higher is the z-score. Losses in the chromosomes show a z-score below minus three and are highlighted in different shades of blue. The lighter the blue colour the smaller is the loss in the chromosomes. The total number of sequencing reads per chromosome arms and the calculated z-scores are provided in the appendix.

Seven out of ten patients show losses of chromosome arm 8p, whereas in nine out of ten patients gains in the chromosome arm 8q were detected. Eight out of ten patients show gains in chromosome arm 5q and losses in chromosome arm 17p. Patient B15 shows the most CNV with seventeen chromosome arms that are affected by gains or losses. Patients B30 and B54 had five chromosome arms which contained gains and losses and were therefore the samples with the lowest number of chromosomes arms affected by CNV.

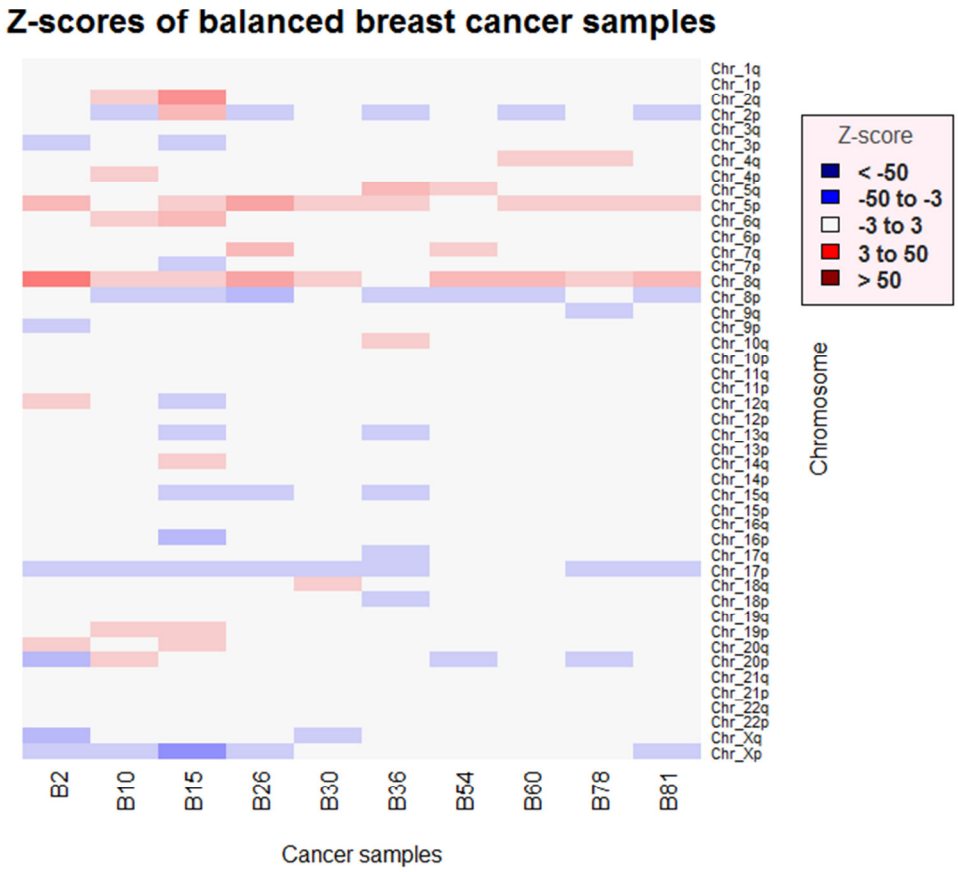


Figure 26: Copy number analysis of breast cancer samples that were classified as balanced by array CGH.

The gains and losses of the chromosome arms were compared with the results of array CGH profiles obtained from a previous study. All array CGH profiles derived from plasma of the breast cancer samples are attached in the appendix. Figure 27 shows the array CGH profile of breast cancer sample B15. A balanced copy number variation profile is indicated by low increase or decrease of the base line at the log2-ratio of zero. Significant gains and losses are marked as green or red parts of the line. FAST-SeqS identified the same chromosomal aberrations as array CGH, if the z-score was significantly higher than three, or lower than minus three. It was not always able to confirm the exact same chromosomal aberrations as array CGH probably due the balanced nature of the genomes and the different basic principles of the used methods. Nevertheless, z-scores derived from FAST-SeqS provide significant results and correlate with the results of the CNV profiles obtained by array CGH.

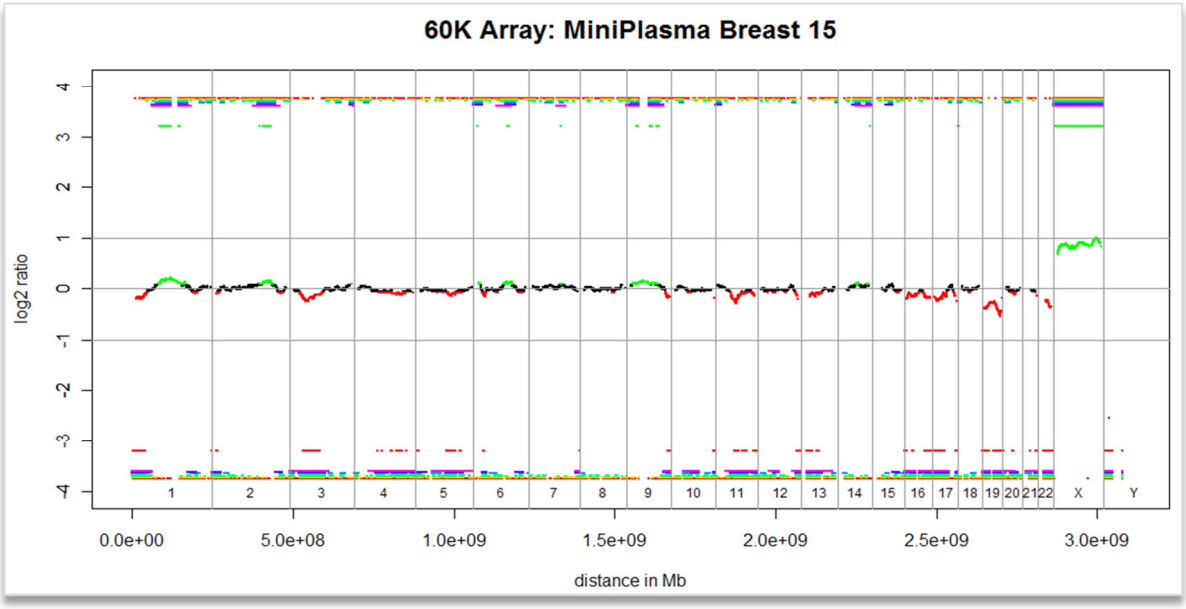


Figure 27: Copy number variation profile of breast cancer sample B15 by array CGH.

Figure 28 shows the z-scores of the chromosome arms of the eighteen breast cancer patients that were diagnosed by array CGH as unbalanced. Z-scores above three signify a gain of the chromosome arm and are emphasized with different shades of red. The darker the colour the higher is the z-score. A dark red colour indicates a z-score above 50. Losses in the chromosomes show a z-score below minus three and are highlighted in different shades of blue. The lighter the blue colour the smaller is the loss in the chromosomes. Dark blue colour emphasizes a z-score that is below minus 50. The total number of sequencing reads per chromosome arms and the calculated z-scores are provided in the appendix.

Seventeen out of eighteen patients show losses of chromosome arm 8p, and in fifteen patients gains in chromosome arm 5q were detected. Most patients also have either gains or losses on the 2p chromosome arm. The majority of patients that were classified previously as unbalanced based on the array CGH profile, because they exhibit a highly altered log2 ratio indicating gains ans losses for some sequences of the genome, showed also lots of gains and losses in the chromosome arms using the Fast sequencing method. In patient B4, B38, B64, B66, and B83 copy number variations were detected for almost all chromosome arms. Still, some patients such as B32 and B 63 had less than ten chromosome arms that showed gains and losses.

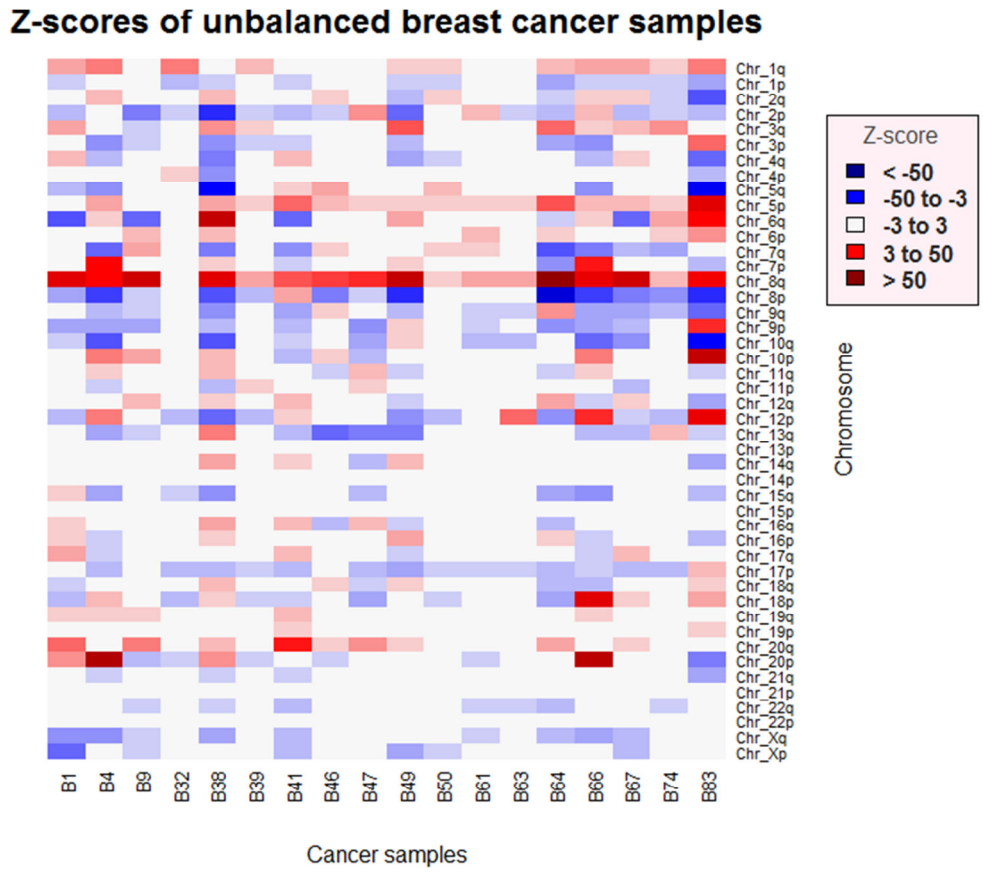


Figure 28: Copy number analysis of breast cancer samples that were classified as unbalanced by array CGH.

The results were compared with the unbalanced array CGH profiles of the breast cancer patients. The array CGH profile of B64 is provided in figure 29. An unbalanced copy number variation profile is indicated by high increase or decrease of the base line at the log2-ratio of zero. Significant gains and losses are marked as red or green parts of the line. The results of FAST-SeqS of B64 were consistent with the results of array CGH.

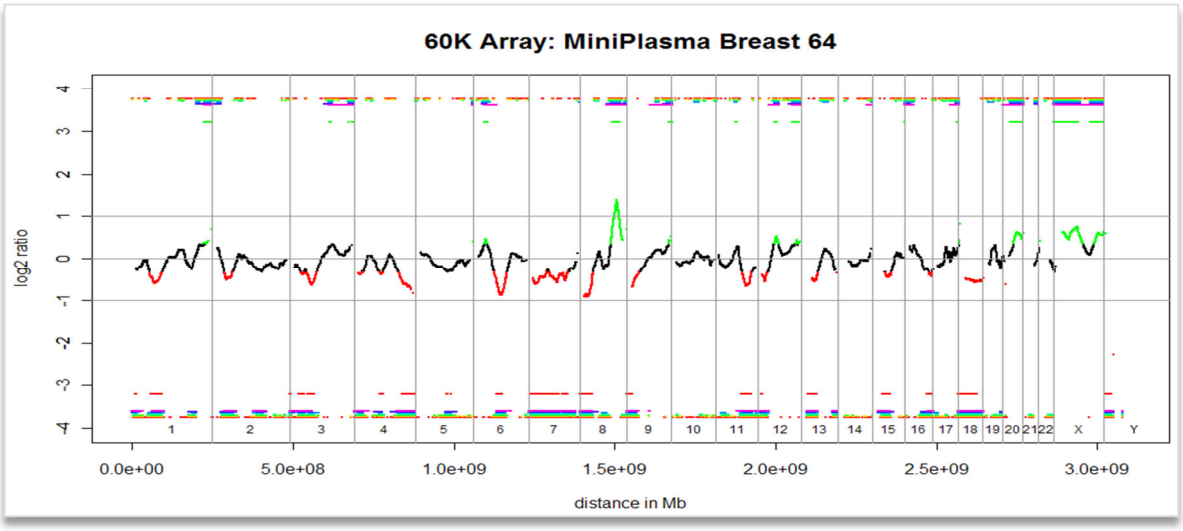


Figure 29: Copy number variation profile of breast cancer sample B64 by array CGH.

FAST-SeqS identified the same chromosomal aberrations as array CGH for all unbalanced samples. Z-scores derived from FAST-SeqS correlate better with the results of the CNV profiles obtained by array CGH if the log2 ratio respectively z-score differs significantly from zero.

The genome wide z-scores of the breast cancer samples were calculated using the z-scores of the chromosome arms and are shown in the table 26. First the sums of square of all chromosome arms were calculated for every breast cancer sample. The sum of square was normalised using the mean and standard deviation of the female control samples.

Table 26: Genome- wide z-score of breast cancer samples. Genome-wide z-scores above 3 are highlighted in red colour.

Breast cancer samples	Sum of Squares	Genome-wide z-scores	Array CGH profile	PlasmaSeq
B1	3205.97	98.22	1	1
B2	511.92	14.09	0	-
B4	6894.65	213.40	1	1
B9	2931.34	89.64	1	1
B10	244,05	5.73	0	-
B15	627.41	17.70	0	1
B26	423.48	11.33	0	-
B30	130.24	2.17	0	0
B32	518.91	14.31	1	1
B36	251.22	5.95	0	-
B38	7753.74	240.23	1	1
B39	501.81	13.78	1	1
B41	2770.71	84.63	1	1
B46	1385.90	41.38	1	1
B47	1792.50	54.08	1	1
B49	4256.23	131.01	1	1
B50	345.62	8.90	1	1
B54	231.50	5.34	0	0
B60	189.30	4.02	0	0
B61	379.43	9.96	1	1
B63	555.05	15.44	1	1
B64	9895.18	307.10	1	1
B66	8384.64	259.93	1	1
B67	3099.99	94.91	1	1
B74	810.32	23.41	1	1
B78	161.88	3.16	0	0
B81	182.53	3.81	0	1
B83	11257.06	349.63	-	-

The table lists also the balanced status for the samples classified previously by array CGH and PlasmaSeq. An unbalanced sample is marked by both techniques with number one and a balanced sample with number zero. A PlasmaSeq copy number variation profile was not available for all breast cancer samples.

Only one of all 28 samples shows a genome-wide z-score lower than three. That sample B30 was classified by the other two methods as balanced. The other samples that were identified by either array CGH or PlasmaSeq as balanced showed a higher genome-wide z-score than three. However, the score remains to be low for those samples. All unbalanced samples were also identified by FAST-SeqS technique with a higher genome-wide z-score than three. Some unbalanced samples had even genome-wide z-score higher than 100 and were consistent with the results of PlasmaSeq and array CGH.

4.3.2 Colon and prostate cancer

Prostate cancer patients were divided into several groups: group 1 is under active surveillance without surgery, group 2 contains plasma samples derived before prostatectomy, group 3 consists of metastatic samples and RTX (radiation treatment) group receives radiation therapy. Patients from group 1 showed a higher prostate-specific antigen (PSA) level in the serum, which is often elevated in presence of prostate cancer. Those patients are therefore frequently screened to prevent overtreatment as not all patients with elevated PSA have prostate cancer and need therapy. The patients of the different groups seem not to be age-matched.

The z-scores of the chromosome arms, short (p) arms and long (q) arms, of prostate cancer patients of group one are provided in the figure 30. Z-scores above three signify a gain of the chromosome arm and are emphasised with different shades of red. The darker the colour the higher is the z-score. Losses in the chromosomes show a z-score below minus three and are highlighted in different shades of blue. The lighter the blue colour the smaller is the loss in the chromosomes. The total number of sequencing reads per chromosome arms and the calculated z-scores are provided in the appendix.

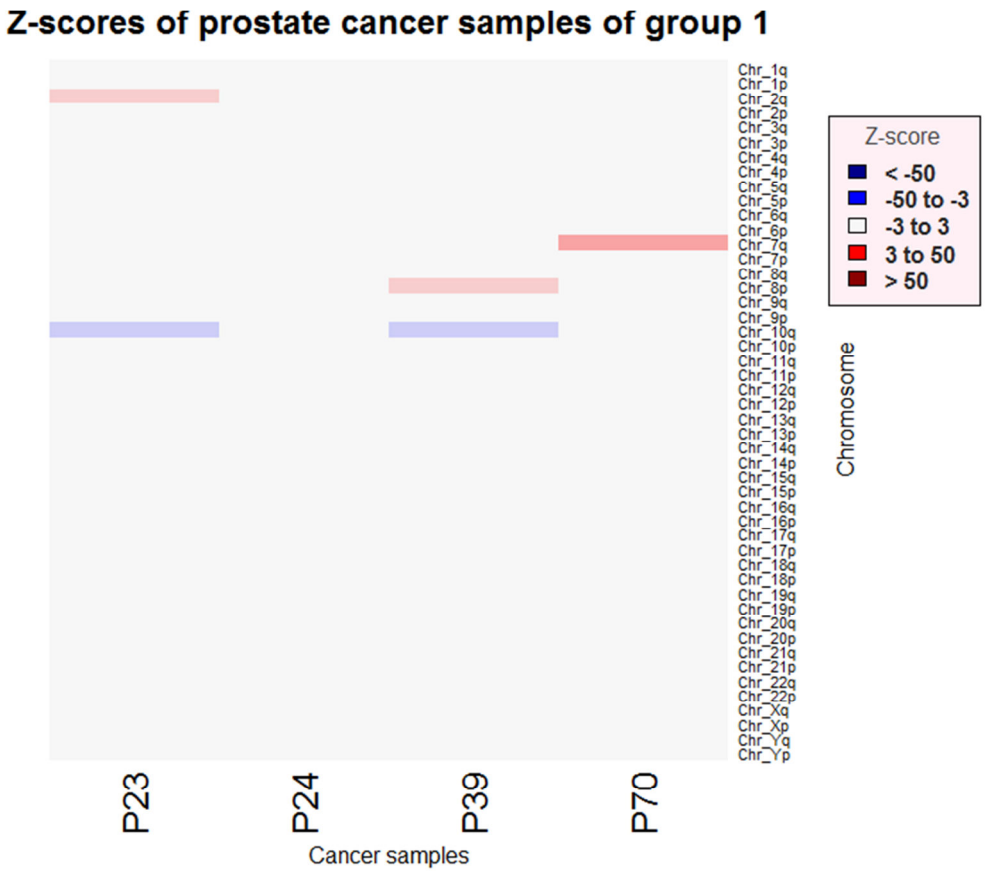


Figure 30: Copy number analysis of prostate cancer samples of group 1.

In figure 30, patient 24 showed no gains or losses of chromosome arms. For patient P70 gains of chromosome arm 7q were observed. In patient P23 gains in chromosome arm 2q and losses in chromosome arm 10q were detected, whereas patient P39 showed also losses in the chromosome arm 10q, but gains in the chromosome arm 8p.

Z-scores of chromosome arms of group two are displayed in figure 31. The results for group two including the total number of sequencing reads per chromosome arms and the calculated z-scores are also added in the appendix.

Three prostate cancer patients of group two (P9, P15 and P16) showed no gains or losses of the chromosome arms. In patient P25 only gains of chromosome arm 10p and in patient P37 gains of chromosome arm 4q were detected. With FAST-SeqS two chromosomal aberrations were detected in patient P71, including gains in chromosome arms 10p and 7q, and three CNV were identified in patient P8 including gains in chromosome arms 7q and 15 p and losses of chromosome arms 16q. Two patients showed significantly more chromosomal aberrations than the others. For patient P81 losses of chromosome arms 3q, 9p, 9q, 10p, and 12q and gains of the chromosome arms 14q and 21p were observed. In patient P8 moderate losses in eight different chromosome arms including 8q, 8p, 12q, 16p 20q, 22p, Yq ,and Yp and high gains in the chromosome X were detected.

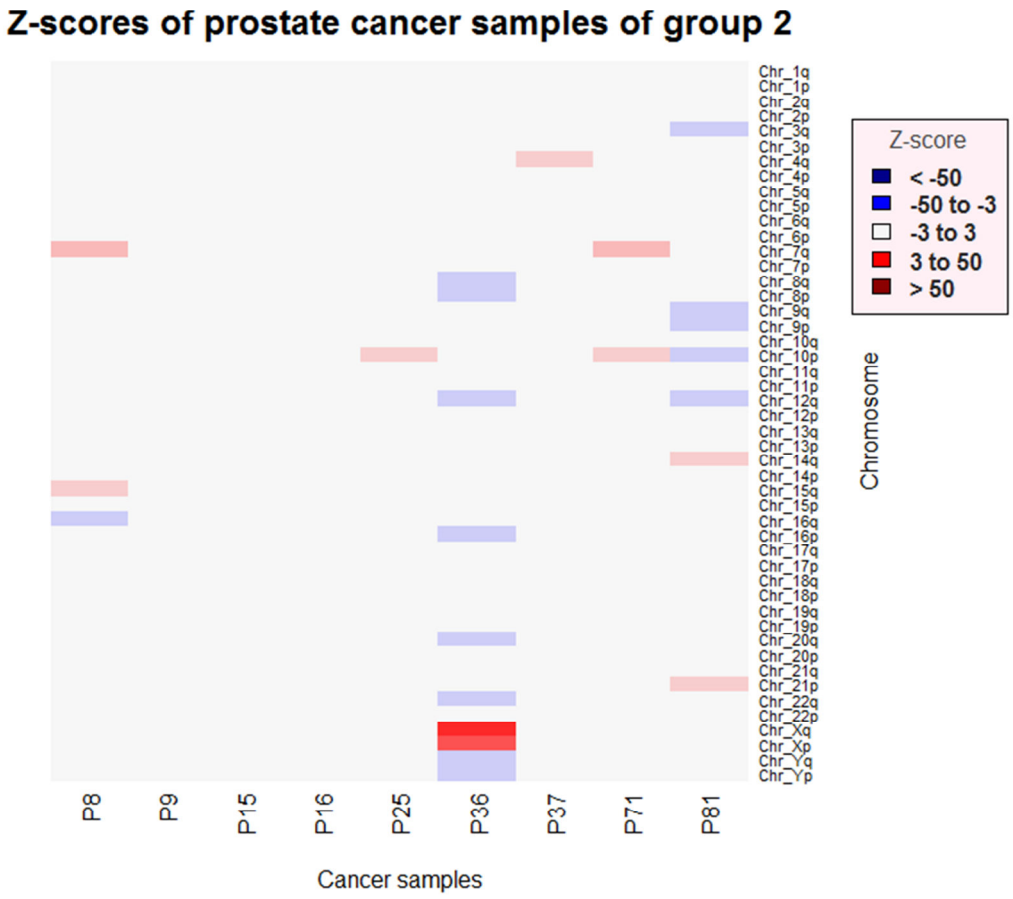


Figure 31: Copy number analysis of prostate cancer samples of group 2.

The metastatic samples of colon and prostate cancer as well as the samples derived of patients receiving radiation therapy showed several gains and losses in different chromosome arms (Figure 32). Some patients like P67, P100-3u, P110o and P127u had had outstanding higher or lower z-scores than the normal range (-3 to +3) and they had therefore the highest numbers of aberrations. Compared to these samples the observed gains and losses for patients P33, P64, P83 and P96 were small including only two to four affected chromosome arms.

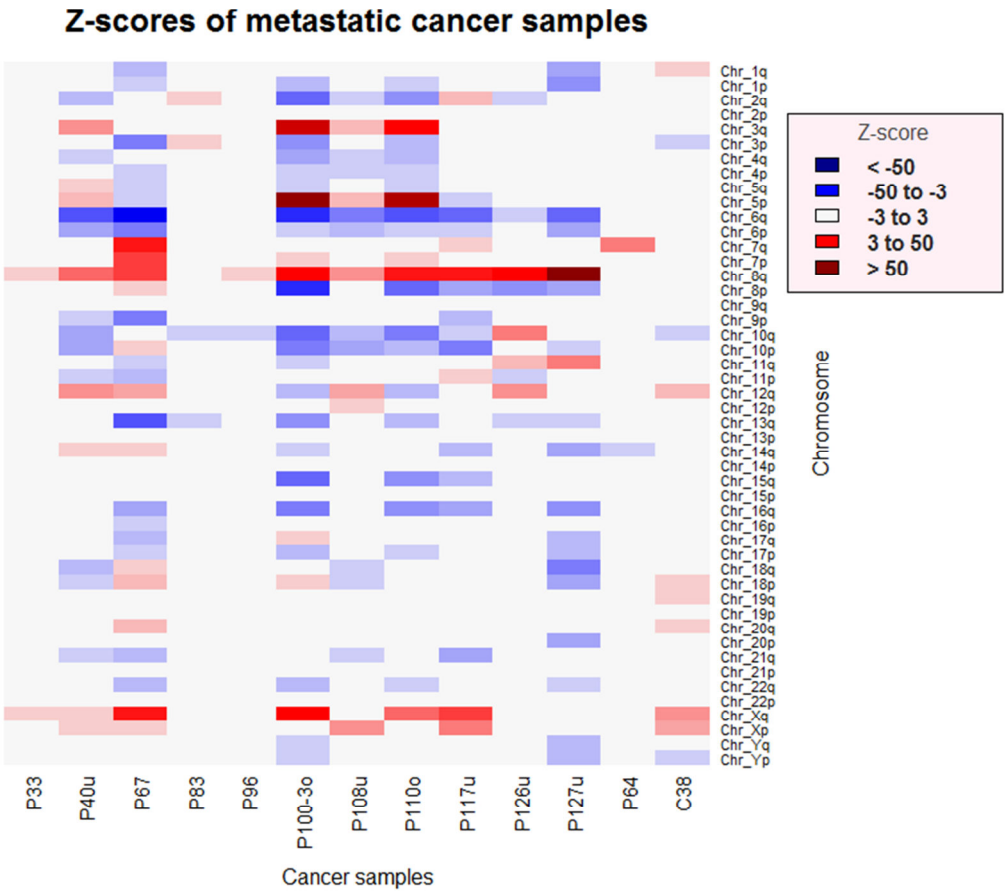


Figure 32: Copy number analysis of metastatic colon and prostate cancer samples.

The gains and losses of the chromosome arms of the third group were compared with the copy number variation profiles of PlasmaSeq obtained from a previous study, because no array CGH was established to date for those samples. PlasmaSeq is a new sequencing approach using a high-throughput sequencing instrument to perform whole genome sequencing from plasma DNA and to generate genome-wide copy number profiles of cancer patients with low cost and time effort. While FAST-SeqS only amplifies a discrete subset of LINE sequences throughout the entire genome, PlasmaSeq enables a whole genome analysis from plasma DNA

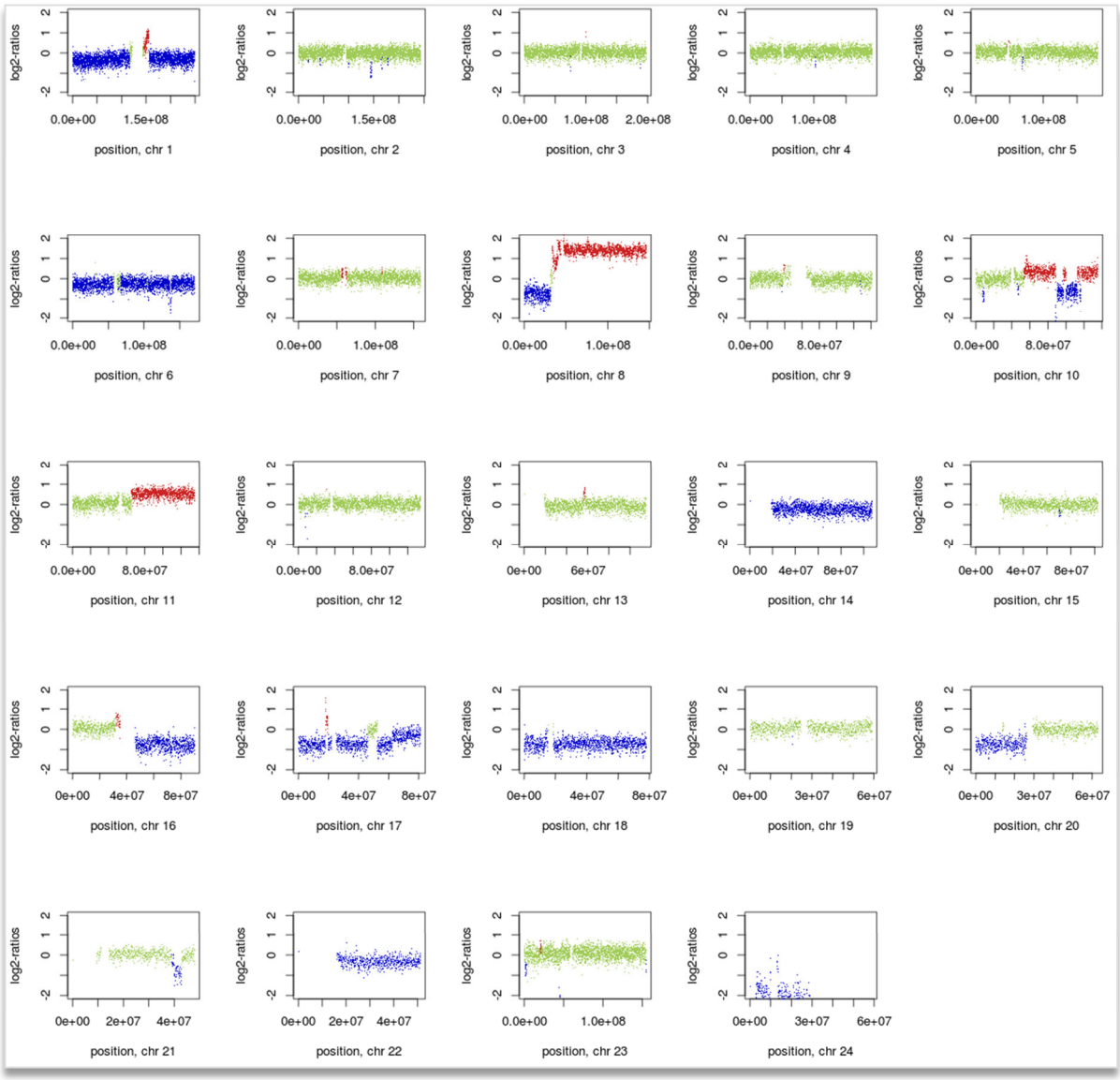


Figure 33: Copy number variation profile of prostate cancer patient P127u.

The PlasmaSeq profiles derived from plasma of the prostate cancer samples are attached in the appendix. Figure 33 shows the PlasmaSeq profile of prostate cancer sample P127u. The copy number variation profile is indicated by an increase or decrease of the base line at the log2-ratio of zero. Significant gains and losses are marked with red or blue. FAST-SeqS identified the almost the same chromosomal aberrations as PlasmaSeq. Both methods detected gains in the chromosomal arms 8p and 11q, losses in the chromosome arms 8p, 10p, 14q, 16q, 20p, and 22q, and losses in the entire chromosomes 1, 6, 17, 18 and Y. FAST-SeqS detected losses in the chromosome arm 13q, which were not detected using PlasmaSeq. On the other hand, PlasmaSeq was able to detect small gains in chromosome 1 and 10 and small losses in chromosome 21.

The low numbers of copy number variations observed with FAST-SeqS were confirmed through comparison with the PlasmaSeq profiles. It was not always able to confirm the same

chromosomal aberrations as with PlasmaSeq due the different basic principals of the used methods. Nevertheless, z-scores derived from FAST-SeqS correlate with the results of the CNV profiles obtained by PlasmaSeq.

The FAST-SeqS results of the colon cancer patient c38 are displayed also in figure 32. There were losses in the chromosome arms 3p, 10q and Yp, and gains in the chromosome arms 1q, 12q, 18p, 19q, 20q. In the entire X chromosome gains were identified. To confirm the detected gains and losses, the sequencing result was compared with a PlasmaSeq profile that was created in a previous study (Figure 34).

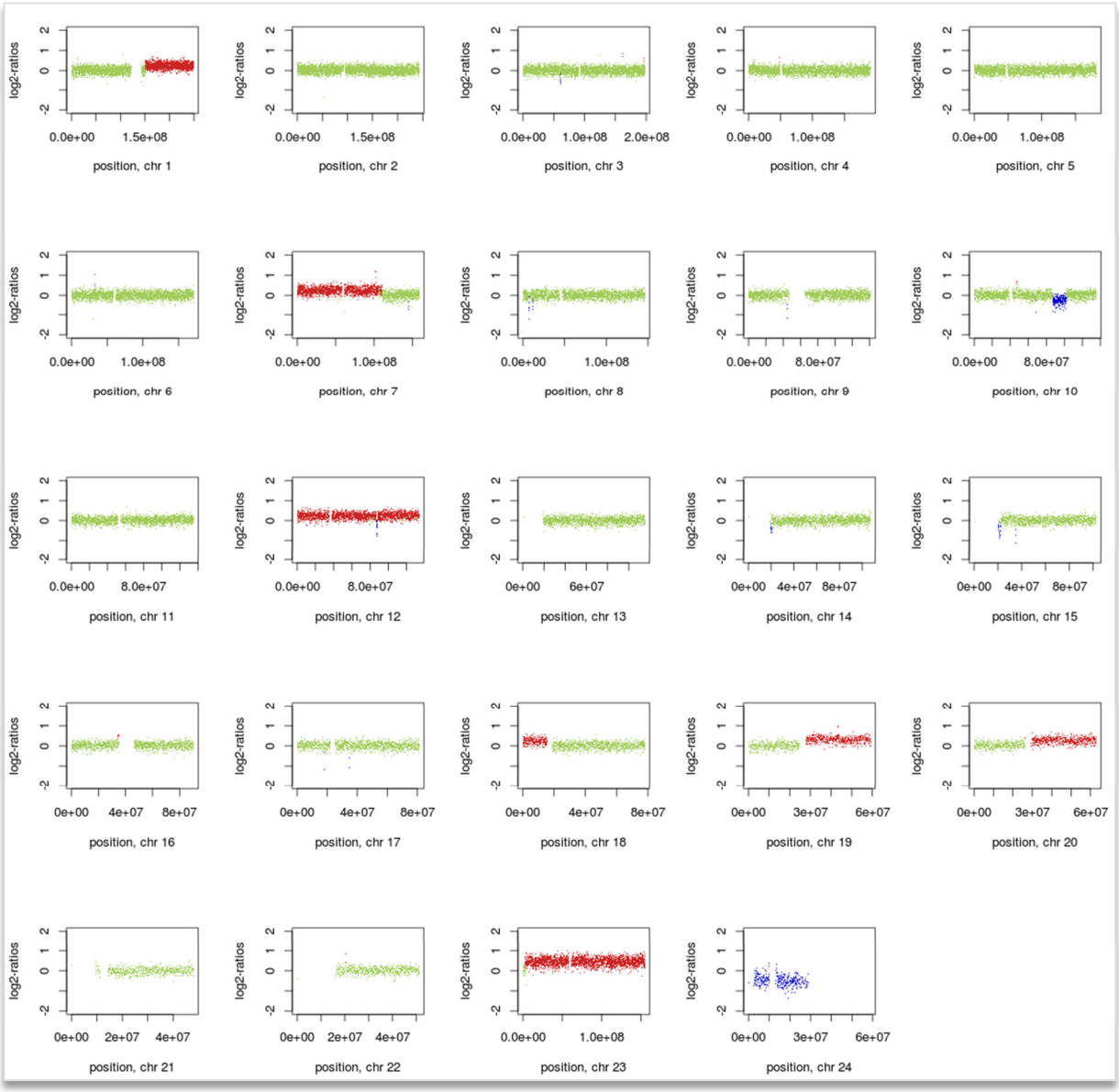


Figure 34: Copy number variation profile of colon cancer patient c38.

The comparison shows a good agreement of the results for both methods. However, FAST-SeqS did not detect the gain in chromosome 7, which was demonstrated by PlasmaSeq. The

moderate losses in the chromosome arm 3p identified by FAST-SeqS were not confirmed by PlasmaSeq.

For the prostate cancer samples as well as for the colon cancer sample, the genome wide z-score was calculated using the z-scores of the chromosome arms. The results are provided in the table 27. The calculated sums of square of every cancer sample were normalised using the mean and standard deviation of the male control samples.

The table assign to each prostate sample its group and the Gleason score, which is commonly used to grade the tumour and predict the prostate cancer outcome. The Gleason score ranges from two to ten, indicating a more aggressive cancer the higher the score becomes. For some cancers, the Gleason score was not calculated yet.

Table 27: Genome-wide z-scores of prostate and colon cancer samples. Genome-wide z-scores above 3 are highlighted in red colour.

Cancer samples	Sum of Squares	Genome-wide z-scores	Group	Gleason score
P8	138.52	2.54	2	6
P9	52.39	-0.07	2	7
P15	37.34	-0.52	2	7
P16	42.56	-0.36	2	7
P23	85.97	0.95	1	6
P24	43.06	-0.35	1	6
P25	58.04	0.10	2	7
P33	109.98	1.68	3	9
P36	1183.95	34.19	2	8
P37	76.55	0.66	2	7
P39	124.39	2.11	1	6
P64	285.53	6.99	RTX	10
P70	161.93	3.25	1	6
P71	114.28	1.81	2	10
P81	146.08	2.77	2	7
P83	101.75	1.43	3	9
P86	18187.78	548.97	2	7
P96	115.54	1.84	3	8
P40u	1582.19	46.25	3	10
P108u	934.42	26.64	3	10
P110o	5814.78	174.38	3	-
P110-3o	11644.29	350.87	3	-
P117u	2375.53	70.26	3	-
P126u	1438.60	41.90	3	-
P127u	12496.06	376.65	3	-
P67	4812.86	144.05	3	10
C38	487.75	13.11		

Three samples of group 1 showed a genome-wide z-score lower than three, whereas patient P70 had a z-score of 3.25 due to the high gains in the chromosome arm 7q. For two patients from group 2, including P36 and P86, a significant high genome-wide z-score was calculated. Only three out of eleven cancer samples of group 3, who have been diagnosed to have metastases, had a genome-wide z-score below three. Samples with high alterations in the chromosome arms have a significantly higher genome-wide z-score that extend to a z-score of 548. For the patient that received radiation therapy a z-score of almost 7 was calculated. In general is the genome-wide z-score higher, if a high grade is determined for the Gleason score.

For the colon cancer sample a genome-wide z-score of 13.11 was calculated and is also displayed in table 27. The high z-score confirms the metastatic state of this tumour.

4.4 DILUTION SERIES OF GENOMIC DNA

DNA from four cancer cell lines were mixed with standard DNA of healthy individuals (Promega) in ratios of 100%, 75%, 50%, 25%, 20%, 15%, 10%, 5%, 2.5%, 1% and 0% in order to evaluate the sensitivity for the detection of chromosomal copy number variations. The z-scores were calculated for each chromosome as well as for each chromosome arm.

4.4.1 WHOLE CHROMOSOME

The sensitivity for the detection of CNV for each chromosome was evaluated for the cancer cell lines HepG2, MCF7 and U2OS. The number of detectable aberrations decreases with the increased amount of DNA derived from normal individuals. The results for the HepG2 dilution series is shown in figure 35. A gain of chromosomal fragments is indicated by a z-score above three and is highlighted with different shades of red. The darker the colour the higher is the z-score. A dark red colour indicates a z-score above 50. Losses in the chromosomes show a z-score below minus three and are highlighted in different shades of blue. The lighter the blue colour the smaller is the loss in the chromosomes. The 2.5% and 1% diluted samples were not prepared and analysed due to lack of cancer cell line DNA.

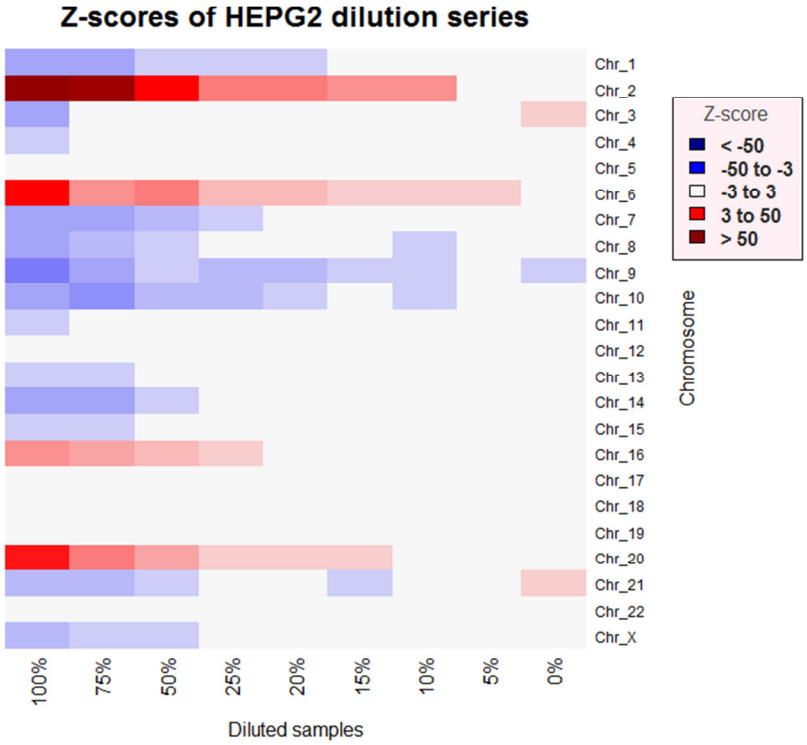


Figure 35: Copy number analysis of cell line HepG2 for each chromosome.

The results of the cancer cell line MCF7 dilution series are provided in the figure 36 and the dilution series of U2OS cancer cell line is shown in figure 37.

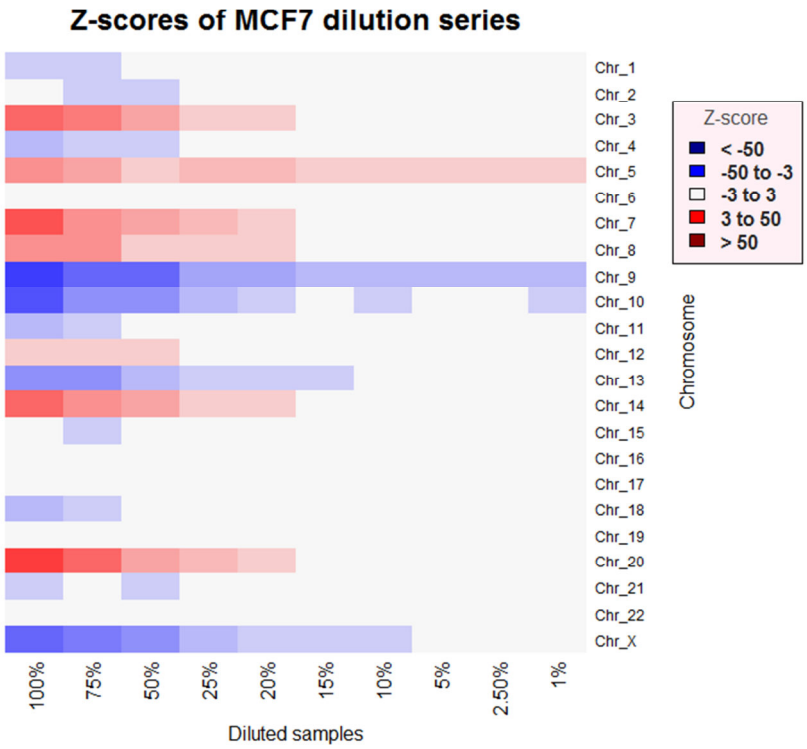


Figure 36: Copy number analysis of cell line MCF7 for each chromosome.

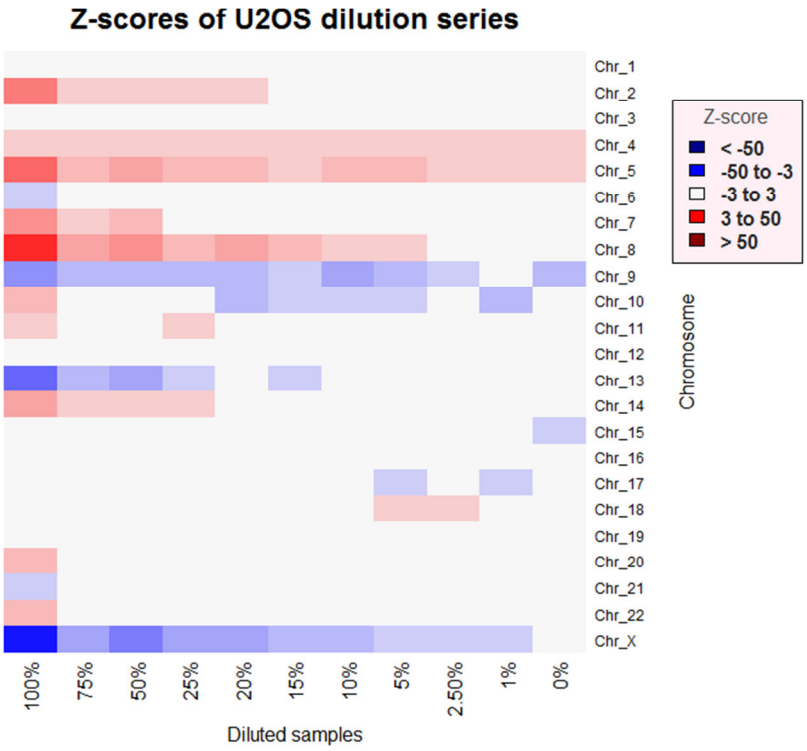


Figure 37: Copy number analysis of cell line U2OS for each chromosome.

As expected all three dilution series shows a decreased number of detected aberrations with the increased amount of DNA derived from normal individuals. Diluted samples with a tumour DNA content below 20 % showed a few gains and losses. Depending on the cell line different chromosome have copy number alterations and suffer from gains and losses to a different extend. Chromosome 2, 8, 13 and X showed great gains and losses in all three

cancer cell lines. The small detected gains and losses in the diluted samples with a tumour DNA content below 10% do not present reliable values that can be used for cancer screening. They might be measured due to dilution and sequencing inaccuracy, as the entire dilution series was sequenced in several runs rather than in one.

4.4.2 Chromosome arm

In addition to the evaluation of the entire chromosomes, analysis of chromosome arms, short (p) arms and long(q) arms, was performed. The short (p) arms of some acrocentric chromosomes including chromosome 13, 14, 15 and 22 were excluded for every dilution series. The proportion of reads presenting each chromosome arm was calculated for each sample by dividing the total read counts that are mapped to that arm by the total number of reads from all chromosomes. This normalised ratio was used to calculate the z-scores of every chromosome arm using the mean and SD of the control samples. The z-scores of the chromosomes arms of the cancer cell line are provided in the figure 38.

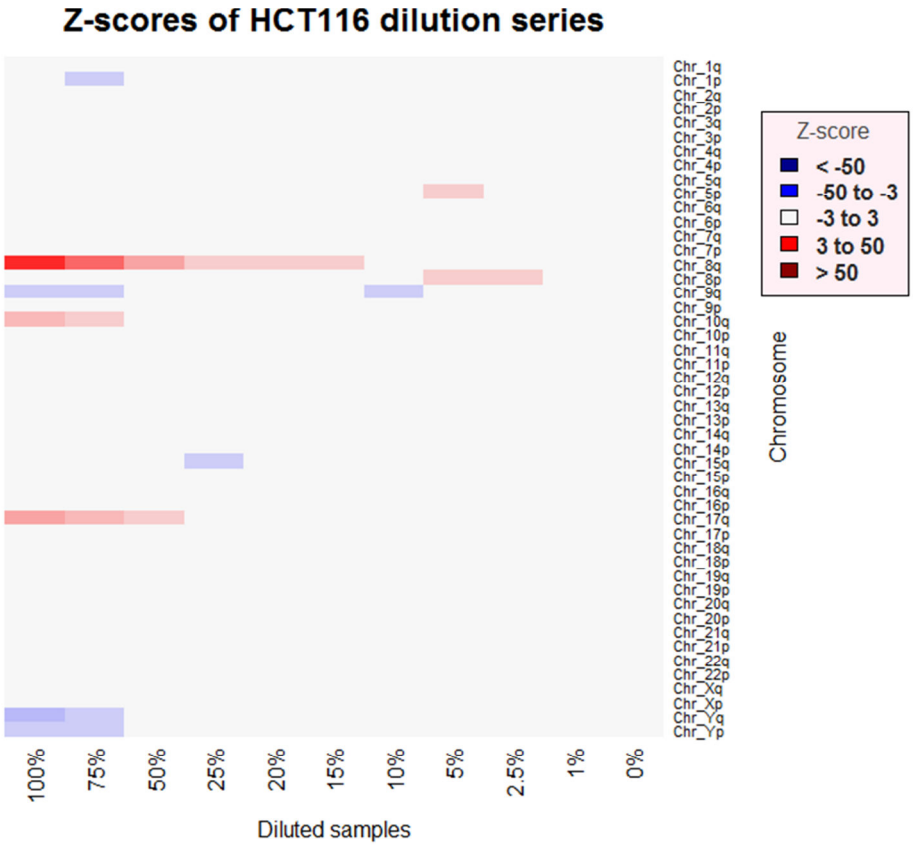


Figure 38: Copy number analysis of cell line HCT116 for each chromosome arm.

A gain of chromosomal fragments is indicated by a z-score above three and is highlighted with different shades of red. The darker the colour the higher is the z-score. Losses in the chromosomes show a z-score below minus three and are highlighted in different shades of blue. The lighter the blue colour the smaller is the loss in the chromosomes.

The results for cancer cell line HepG2 is displayed in figure 39. The 2.5% and 1% diluted samples were not prepared and analysed due to lack of HepG2 DNA. The z-scores of cell line MCF7 are provided in figure 40 and the results displayed in figure 41 show the z-scores for the dilution series of U2OS cell line.

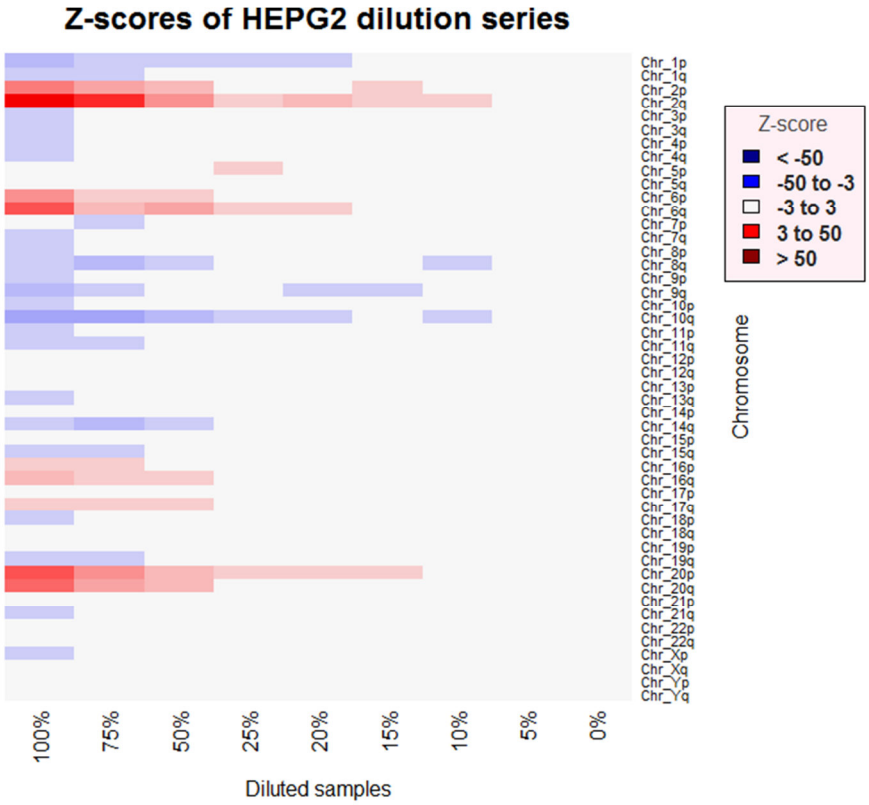


Figure 39: Copy number analysis of cell line HepG2 for each chromosome arm.

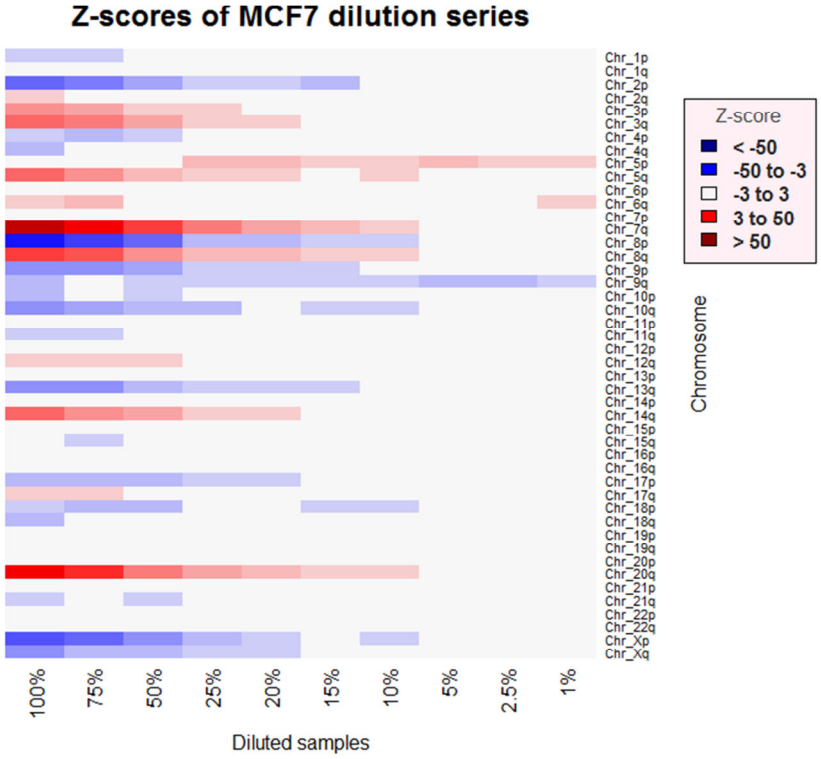


Figure 40: Copy number analysis of cell line MCF7 for each chromosome arm.

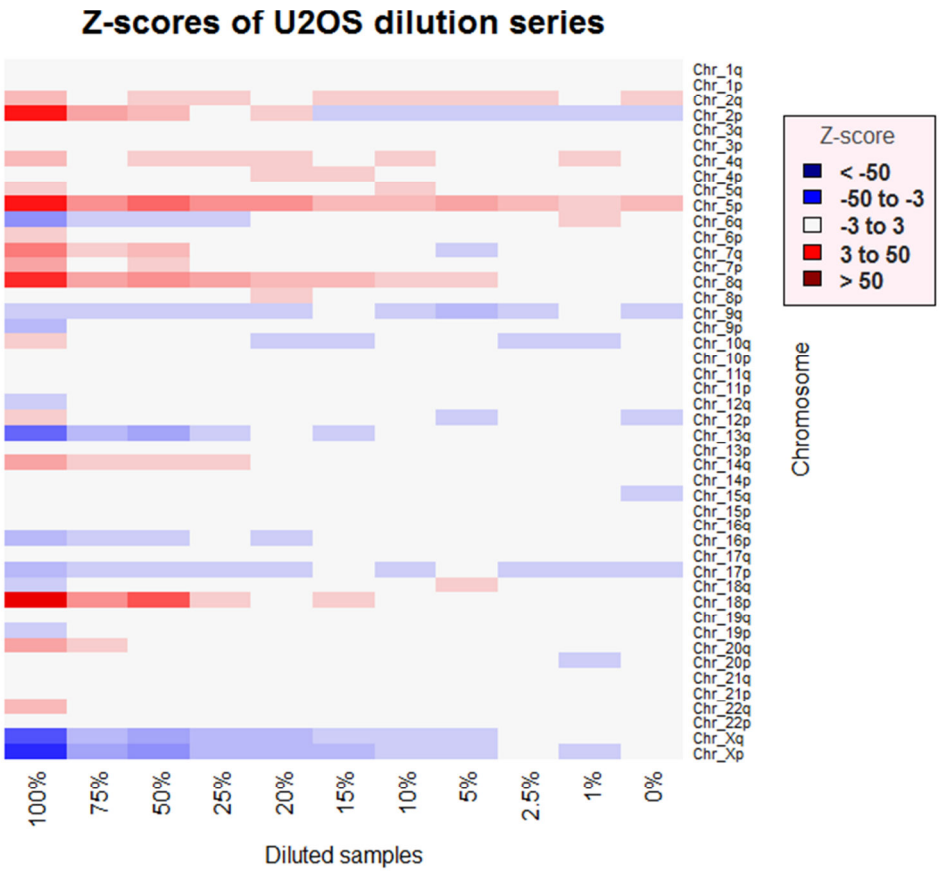


Figure 41: Copy number analysis of cell line U2OS for each chromosome arm.

As expected decreases the number of detected chromosomal aberrations with increased amount of healthy individual's DNA. Diluted samples with a tumour DNA content below 15 % showed a few gains and losses.

Cancer cell line HCT 116 is a colon colorectal carcinoma with only a few copy number alterations in the diluted samples with a tumour DNA content above 15 % including significant gains in chromosome arm 8q and small gains in chromosome arm 17q and 10 q and small losses in chromosome Y. In contrast, the other three cancer cell lines harbour numerous gains and losses in chromosome arms. The liver carcinoma HepG2 shows significant gains in the entire chromosome 2, 6 and 20 and smaller gains in chromosome 16 and in chromosome arm 17q. Chromosome arms of HepG2 that are affected by losses are for example 10q, 14q and 8q, which can be used together with the observed gains to detect cancer. The breast cancer cell line MCF7 has high gains in chromosome arms 3q, 3p, 5q, 7q, 8q, 14q and 20q and great losses in the chromosome arms 2p, 8p, 9p, 10q, 13q,Xp and Xq. Whereas, the osteosarcoma cell line U2OS shows extreme losses in chromosome arms 13q and 6q and in the entire X chromosome and great gains in the chromosome arms 2p, 5p, 7q, 8q, 14q and 18p. Those z-scores of the chromosome arms that are highly altered could be used for cancer detection, because most of them were still detectable at a tumour DNA concentration of 15 %.

A genome-wide z-score was calculated for all four cancer cell lines using the z-scores of the chromosome arms, because they showed a more reliable characteristic of dilution. The results that are provided in the table 28 are highlighted red if they have a higher genome-wide z-score than three. First the sums of square of all chromosome arms were calculated for every cell line. The sum of square was normalised using the mean and standard deviation based on the chromosome arm reads of the control samples of the same gender. All four cell line showed a high genome-wide z-score for tumour DNA content above 75%. Genome-wide z-scores of cell lines MCF7 and HepG2 present a good dilution series with a constantly declining score. Genome-wide z-scores of HCT116 dilution series above three were only calculated for tumour DNA content above 75% due to its few copy number alterations, whereas all diluted samples of the cell line U2OS showed a genome-wide z-score above three.

Table 28: Genome-wide z-scores of dilution series of cell lines. Genome-wide z-scores above 3 are highlighted in red colour.

Cell line	Sum of Squares	Genome-wide z-scores	Cell line	Sum of Squares	Genome-wide z-scores
MCF7 100%	7003.16	216.79	**HepG2 100%**	2834.58	84.16
MCF7 75%	4207.40	129.49	**HepG2 75%**	1286.34	37.29
MCF7 50%	2789.34	85.21	**HepG2 50%**	568.99	15.57
MCF7 25%	691.77	19.71	**HepG2 25%**	204.79	4.55
MCF7 20%	456.78	12.37	**HepG2 20%**	153.98	3.01
MCF7 15%	303.02	7.57	**HepG2 15%**	132.95	2.37
MCF7 10%	240.53	5.62	**HepG2 10%**	124.78	2.12
MCF7 5%	154.42	2.93	**HepG2 5%**	60.17	0.17
MCF7 2.5%	132.74	2.25	**HepG 2.5%**	-	-
MCF7 1%	117.53	1.78	**HepG2 1%**	-	-
MCF7 0%	-	-	**HepG2 0%**	53.02	-0.05
U2OS 100%	5302.41	163.68	**HCT116 100%**	790.31	22.27
U2OS 75%	814.79	23.55	**HCT116 75%**	450.79	11.99
U2OS 50%	1351.60	40.31	**HCT116 50%**	152.83	2.97
U2OS 25%	459.74	12.46	**HCT116 25%**	85.36	0.93
U2OS 20%	509.85	14.03	**HCT116 20%**	65.08	0.32
U2OS 15%	288.63	7.12	**HCT116 15%**	60.91	0.19
U2OS 10%	251.40	5.96	**HCT116 10%**	58.36	0.11
U2OS 5%	334.86	8.56	**HCT116 5%**	73.13	0.56
U2OS 2.5%	205.51	4.52	**HCT116 2.5%**	65.11	1.63
U2OS 1%	185.71	3.91	**HCT116 1%**	56.36	0.05
U2OS 0%	219.67	4.97	**HCT116 0%**	39.91	-0.44

4.5 Dilution series of fragmented genomic DNA

To simulate plasma DNA the genomic DNA of all four cancer cell lines was fragmented to an average size of 250 bp. Then, the fragmented DNA was mixed with fragmented standard DNA of healthy individuals in ratios of 100%, 75%, 50%, 25%, 20%, 15%, 10%, 5%, 2.5%, 1% and 0% in order to evaluate the sensitivity for the detection of chromosomal copy number variations using short fragmented DNA. The z-scores were calculated for each entire chromosome as well as for each chromosome arm.

4.5.1 Whole chromosome

The sensitivity for the detection of CNV for each chromosome was evaluated only for the fragmented DNA derived of the MCF7 cancer cell line. The number of detectable aberrations decreases with the increased amount of DNA derived from normal individuals. The results for the MCF7 dilution series is provided in figure 42. A gain of chromosomal fragments is highlighted with different shades of red indicating a z-score above three. The darker the colour the higher is the z-score. Losses in the chromosomes show a z-score below minus three and are highlighted in different shades of blue. The lighter the blue colour the smaller is the loss in the chromosomes. The pure DNA of healthy individuals (0%) was not sequenced and analysed. As the amount of fragmented MCF7 cell line DNA decreases the z-score declines similar to the results with non-fragmented DNA. The fragmented dilution series showed similar gains and losses as the non-fragmented MCF7 dilution series.

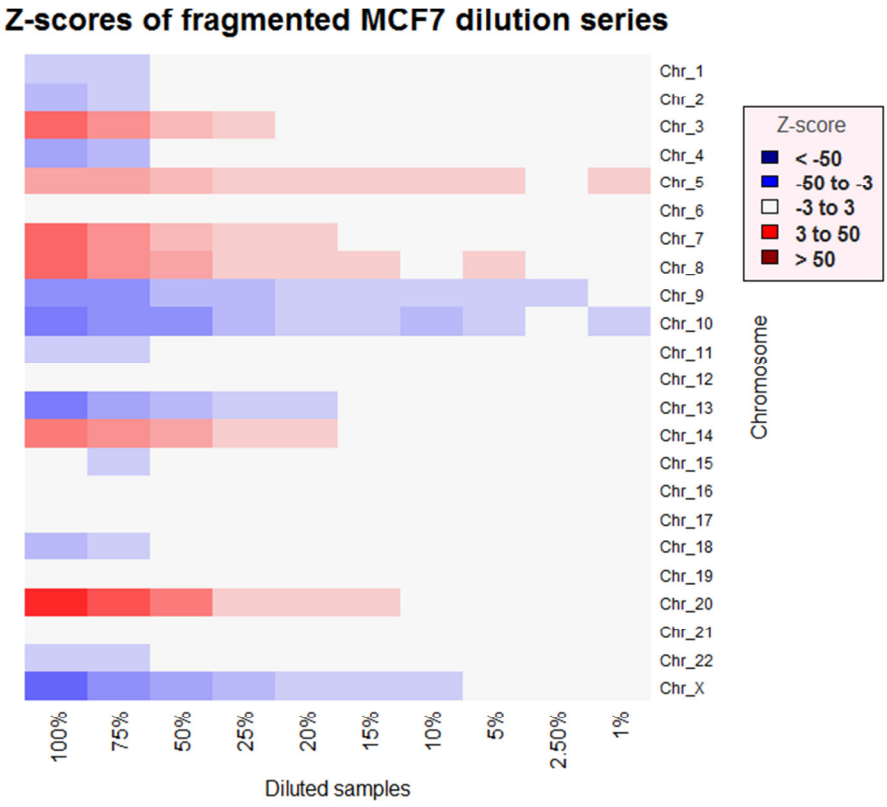


Figure 42: Copy number analysis of fragmented cell line MCF7 for each chromosome.

4.5.2 Chromosome arm

For the evaluation of the entire chromosomes, analysis of chromosome arms was performed also for the fragmented dilution series. Data filtration and analysis were conducted in the same way as for the non-fragmented dilution series. After excluding the short (p) arms of chromosomes 13, 14, 15 and 22, and normalisation, the normalised ratio was used to calculate the z-scores of every chromosome arm using the mean and SD of the control samples. The z-scores of the autosome arms of the fragmented HCT116 cell line are provided in the figure 43 that displays only a significant gain in chrosmosome arm 8q. The results for all autosomes of the copy number analysis of the fragmented HepG2 cell line are presented in figure 43. The fragmented HepG2 cell lines harbours the same copy number alterations as the non-fragmented cell line including gains in the entire chromosome 2, 6, 16 and 20 and losses in chromosomes such as 1, 8, 10 and 14. The z-scores based on the chromosome arms of the fragmented cancer cell line MCF7 are shown in figure 45 including significant gains in chromosome arms 3p, 3q, 5q, 7q, 8p, 14q and 20q and great losses in chromosome arms 2p, 8p, 9p, 10q, 13q and Xp, and the calculated results of the fragmented cancer cell line U2OS are provided in figure 46. The fragmented U2OS cell line has high gains in chromosome arms 2p, 5p, 7q, 8q and 18p and high losses in chromosome arms 6q, 13q, Xp and Xq.

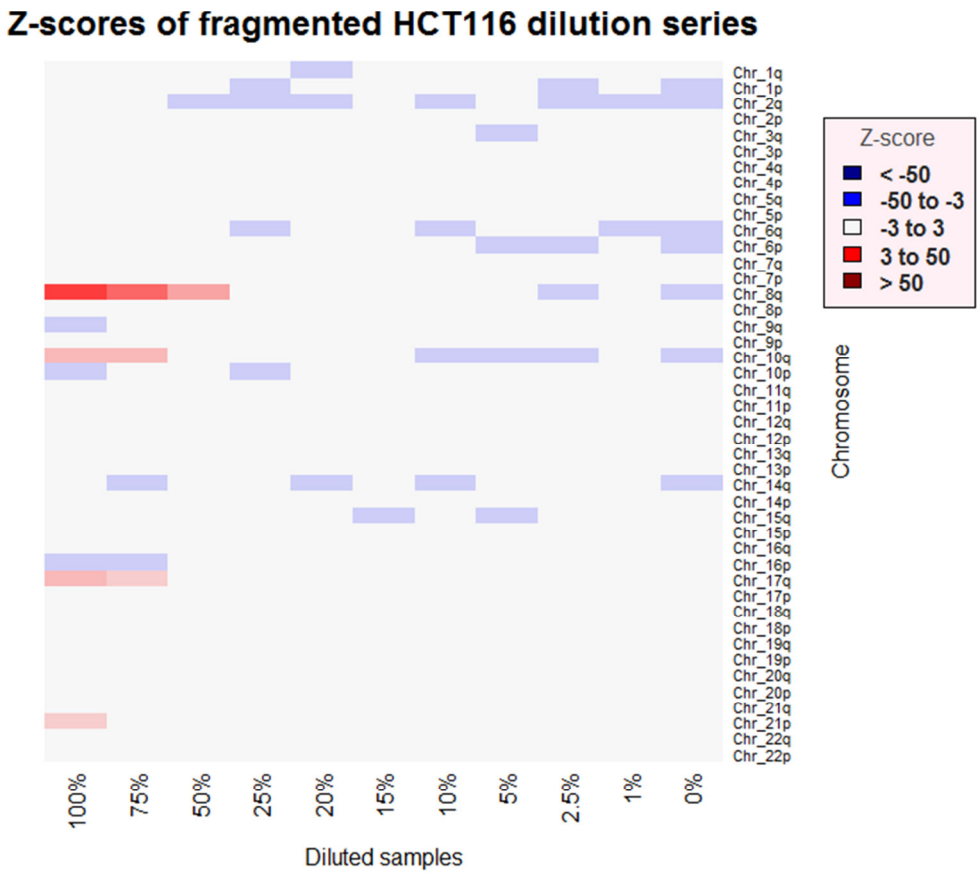


Figure 43: Copy number analysis of fragmented cell line HCT116 for each autosome arm.

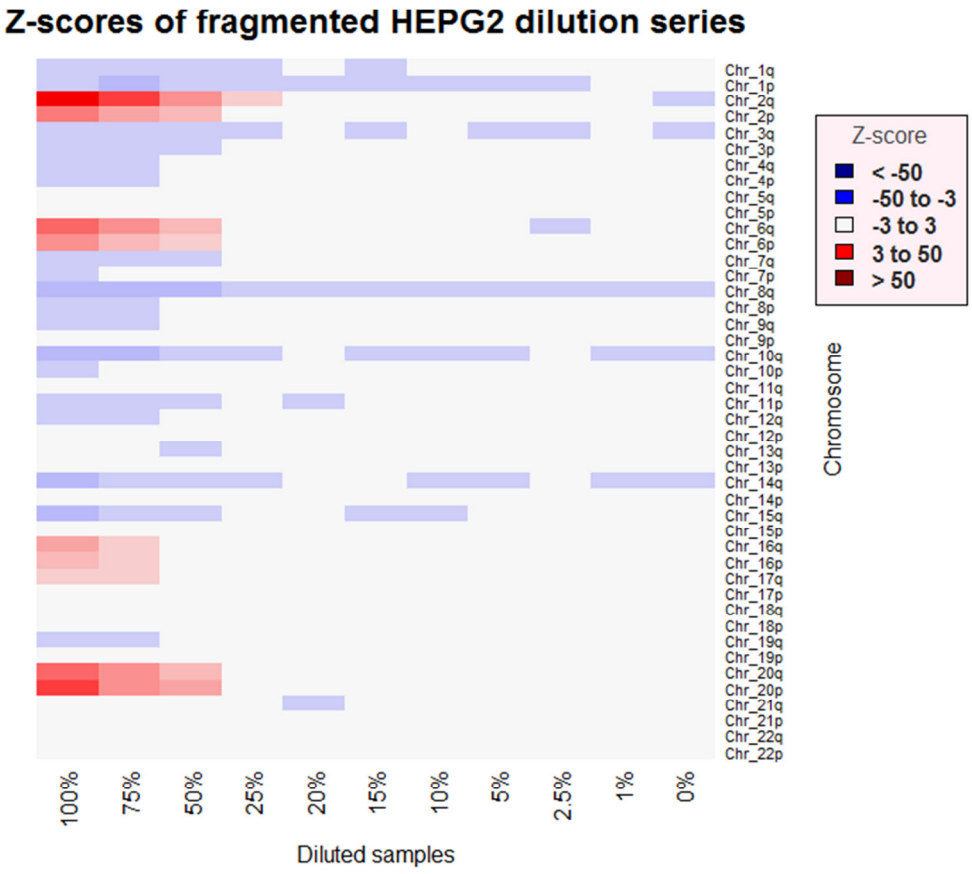


Figure 44: Copy number analysis of fragmented cell line HepG2 for each autosome arm.

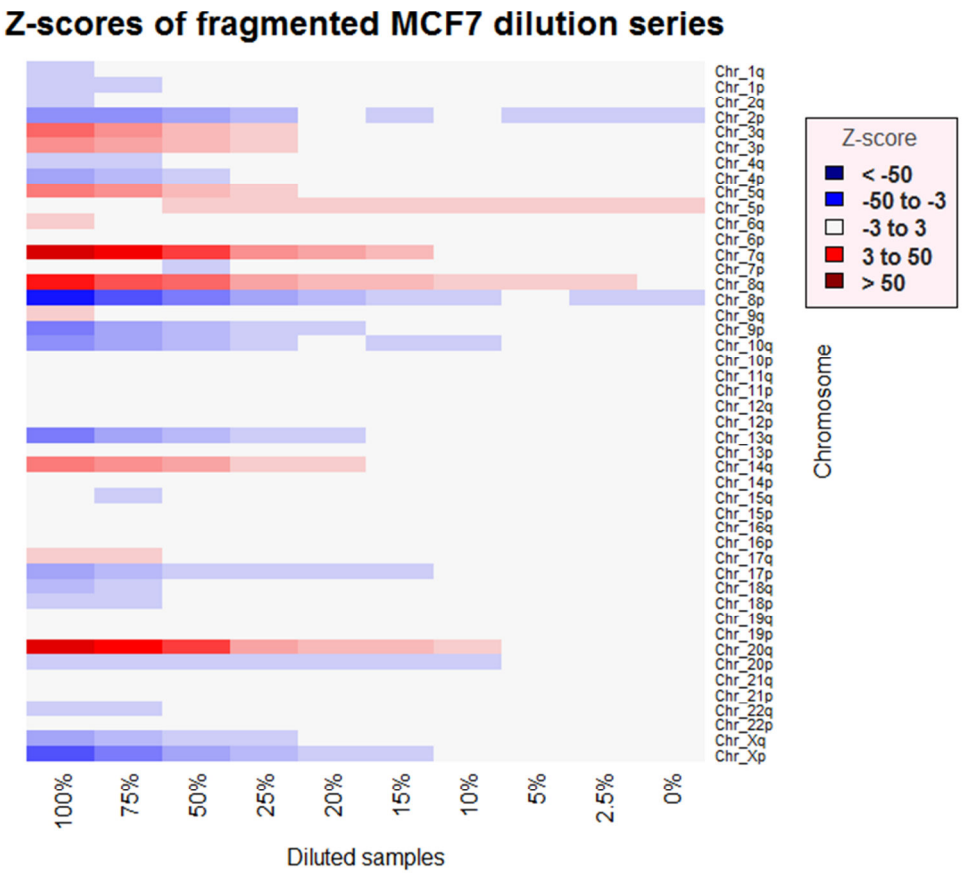


Figure 45: Copy number analysis of fragmented cell line MCF7 for each chromosome arm.

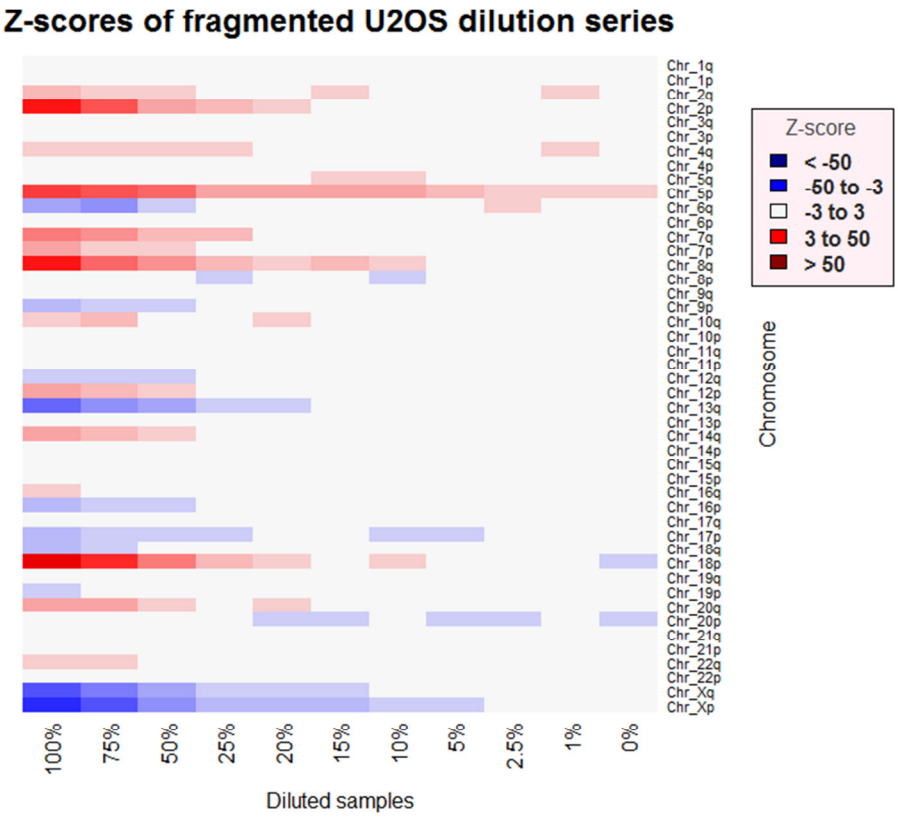


Figure 46: Copy number analysis of fragmented cell line U2OS for each chromosome arm.

As expected the number of detected chromosomal aberrations decreases with increased amount of healthy individual's DNA. Diluted samples with a tumour DNA content below 15 % showed a few gains and losses for all four cell lines. The z-scores of the fragmented cancer cell line HCT116 incline almost not with an increased amount of Promega DNA and no dilution effect can be observed except for chromosome arms 8q and 17q.

A genome-wide z-score was calculated for all four cancer cell lines using the z-scores of the chromosome arms, because they showed a more reliably characteristics of dilution. The results for the cell lines MCF7 and U2OS are provided in the table 29, whereas the autosome-wide z-scores of cell lines HepG2 and HCT116 are listed in the table 30. Z-scores that are above three are highlighted red. The genome-wide z-score was calculated for the fragmented cell lines MCF7 and U2OS by standardizing the normalised sums of square of all chromosome arms using the mean and standard deviation of the female control samples. The reads located on the X- and Y-chromosomes of the cell lines HepG2 and HT116 were omitted, because the cancer cell line DNA was mixed with Promega DNA of the opposite gender. The sum of squares was calculated for all autosomes and then standardized using the mean and standard deviation of the male control samples. All four fragmented cell lines showed a high z-score for tumour DNA content above 25%. Furthermore, all diluted samples of the fragmented cell line MCF7 showed a genome-wide z-score above three. The 1% diluted sample has a too high z-score and is not reliable; the rest of the dilution series is constantly declining. Genome-wide z-scores of fragmented cell line U2OS present a good

dilution series with a constantly decreasing score except the 0% diluted sample, which shows a higher genome-wide z-score than the 1% sample.

Table 29: Genome-wide z-scores of fragmented cell lines MCF7 and U2OS. Genome-wide z-scores above 3 are highlighted in red colour.

Cell line	Sum of Squares	Genome-wide z-scores	Cell line	Sum of Squares	Genome-wide z-scores
MCF7 100%	6715.94	207.82	**U2OS 100%**	5275.33	162.84
MCF7 75%	4061.41	124.93	**U2OS 75%**	2927.45	89.52
MCF7 50%	2038.94	61.78	**U2OS 50%**	1270.51	37.78
MCF7 25%	705.31	20.13	**U2OS 25%**	462.60	12.55
MCF7 20%	497.60	13.65	**U2OS 20%**	324.79	8.25
MCF7 15%	382.85	10.06	**U2OS 15%**	511.35	14.07
MCF7 10%	274.42	6.68	**U2OS 10%**	351.56	9.08
MCF7 5%	192.46	4.12	**U2OS 5%**	210.47	4.68
MCF7 2.5%	200.11	4.36	**U2OS 2.5%**	163.04	3.20
MCF7 1%	41157.32	1283.30	**U2OS 1%**	115.36	1.71
MCF7 0%	162.73	3.19	**U2OS 0%**	129.90	2.16

Autosomal-wide z-scores of fragmented HCT116 dilution series above three were only calculated for tumour DNA content above 25% and for the 0% sample. The autosomal-wide z-scores of both fragmented cell lines are similar to the genome-wide z-scores of the same non-fragmented cell lines. The fragmented cell line HepG2 shows an inclining detectable autosomal-wide z-score including the 15% diluted sample with the exception of the 20% sample, which has a score lower than three. The 0% diluted samples of both fragmented cell lines have an autosomal-wide z-score slightly higher than three.

Table 30: Autosome-wide z-scores of fragmented cell lines HepG2 and HCT116. Genome-wide z-scores above 3 are highlighted in red colour.

Cell line	Sum of Squares	Genome-wide z-scores	Cell line	Sum of Squares	Genome-wide z-scores
HepG2 100%	2864.73	85.07	**HCT116 100%**	674.58	18.77
HepG2 75%	1482.90	43.24	**HCT116 75%**	380.10	9.85
HepG2 50%	637.02	17.63	**HCT116 50%**	166.11	3.38
HepG2 25%	207.93	4.64	**HCT116 25%**	160.23	3.20
HepG2 20%	149.00	2.86	**HCT116 20%**	123.95	2.10
HepG2 15%	164.98	3.34	**HCT116 15%**	108.00	1.62
HepG2 10%	144.66	2.73	**HCT116 10%**	144.79	2.73
HepG2 5%	151.46	2.93	**HCT116 5%**	142.75	2.67
HepG2 2.5%	141.04	2.62	**HCT116 2.5%**	142.36	2.66
HepG2 1%	149.90	2.89	**HCT116 1%**	144.31	2.72
HepG2 0%	154.79	3.03	**HCT116 0%**	161.18	3.23

The number of detected chromosomal aberrations decreases with increased amount of healthy individual's DNA for each fragmented and non-fragmented cancer cell lines. The fragmented cancer cell lines showed the same gains and losses as the non- fragmented DNA of the same cell line. Diluted samples with a tumour DNA content above 20 % could be used for cancer screening using specific alteration in the chromosome arms or entire chromosomes. Genome-wide z-scores also commonly decreased with smaller amount of tumour DNA, whereas the detection limit using a z-score of 3 differes among the cancer cell lines. HepG2 and HCT116 have a lower genome-wide z-score than MCF7 and U2OS. Therefore, the sensitivity of the genome-wide z-score needs to be further analysed in order to use it for copy number alteration detection.

4.6 WHOLE- GENOME LIBRARIES OF CELL LINES

The results from the fast library preparation method should be compared with an established whole-genome library preparation technique. Therefore, a common used whole-genome amplification (WGA) kit from Illumina was used to examine the gains and losses of the four cell lines.

The results of the copy number variation profiles could not be normalised due to the lack of control samples. The profiles were generated using the read counts mapped to the reference genome. The CNV profile of cell line HCT116 is provided in figure 47.

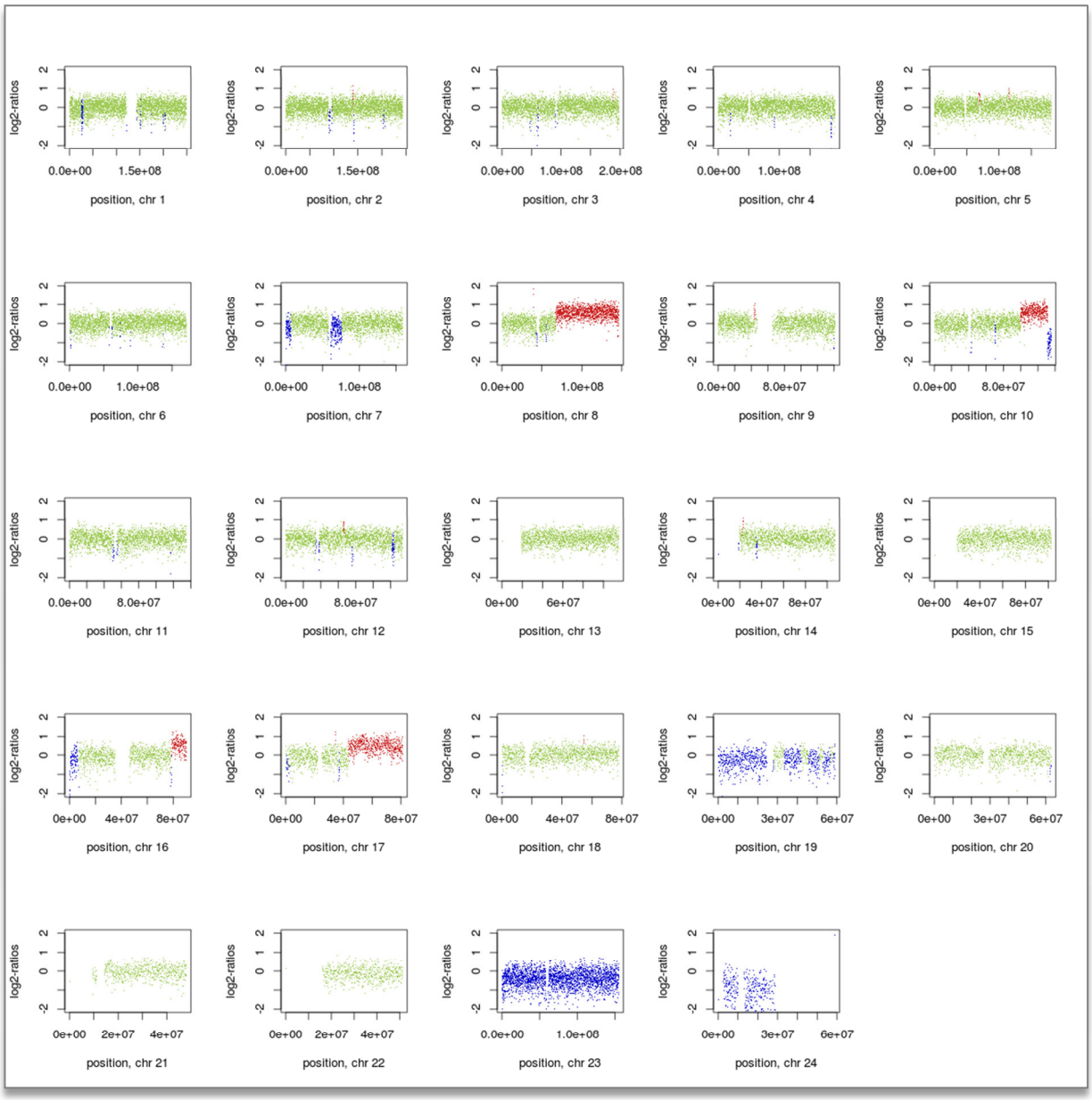


Figure 47: Copy number variation (CNV) profile of cell line HCT116.

FAST-SeqS as well as the WGA method detected losses in the chromosome Y and in the chromosome arm 16p, and gains in the chromosome arms 8q, 10q and 17q. FAST-SeqS also detected in the non-fragmented plasma samples losses in chromosome arm 9q that were not confirmed by the WGA profile. In addition, the CNV profiles obtained from the WGA library shows losses in the chromosomes 7, 10, 19 and X, which were not identified by FAST-SeqS using either fragmented or non-fragmented cell line.

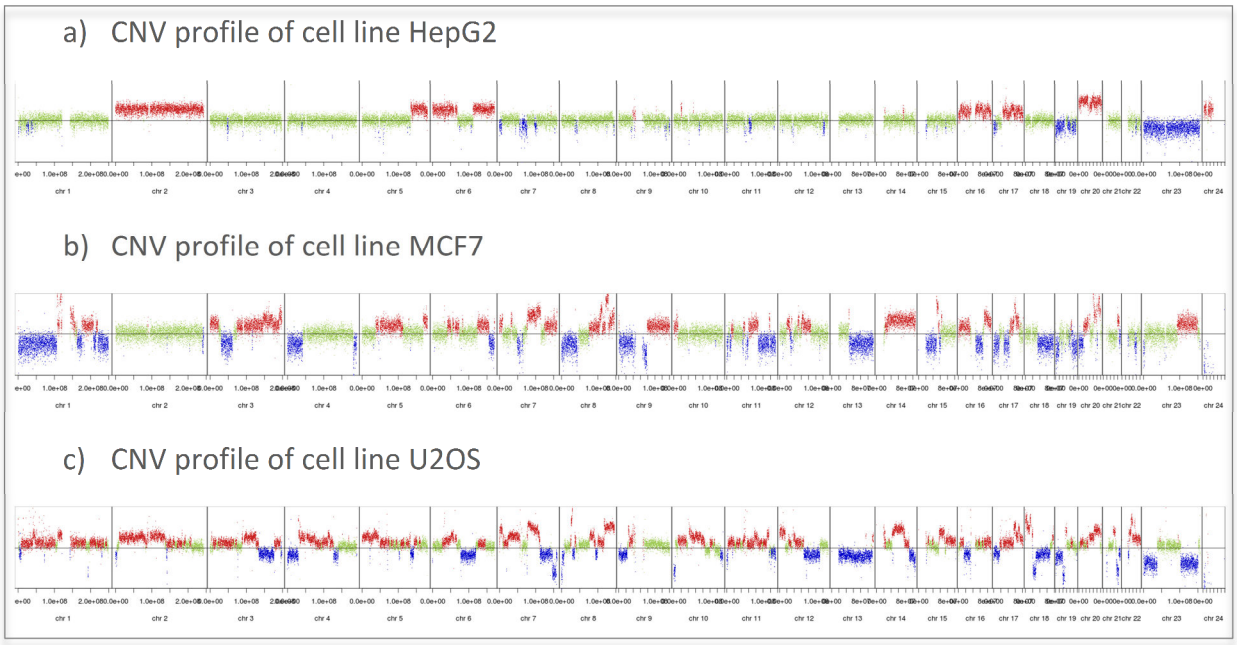


Figure 48: Copy number variation (CNV) profiles of cancer cell lines.

The profiles of the remaining three cancer cell lines are presented in figure 48. However, it was not possible to create good and reliable CNV profiles for the other three cell lines without control samples. The profiles showed some similarities between the gains and losses compared to the FAST-SeqS results such as gains in the chromosome 2, 6 and 20 in HepG2 cancer cell line. The highly altered copy number alterations in cell lines MCF7 and U2OS support the results that were obtained with FAST-SeqS method, but for a better and more accurate comparison the CNV profiles need to be normalised with control samples.

5. DISCUSSION

Development and access to accurate and sensitive screening methods for the early and fast detection of cancer and identification of its related treatment is of utmost importance. FAST-SeqS was established recently as a safe and effective screening method for fetal aneuploidy achieving increased throughput and decreased cost. Kinde *et al.* (2012) reported with their improved approach a method that enables to monitor up to 100 DNA samples in a few days for aneuploidy, which is also a common characteristic of tumours manifesting at the earliest stage of tumour development.

Recently, analysis of chromosomal alterations in tumour such as abnormal chromosomal copy number changes as well as rearrangements enabled the development of patient-specific biomarkers. The complex nature of cancer emphasised the need for personalised medicine approaches including prognosis and diagnostics. Research work had identified biomarkers, which have been realised in clinic, at several levels: prognostic, predictive and pharmacokinetic biomarkers [Heitzer *et al*. (2013a)]. The frequent lack of biopsiable tumour tissue makes blood-based assays crucially important. Non-invasive methods such as circulating tumour cells, plasma DNA, or mRNA and microRNA expression signatures can be used to monitor cancers. Especially plasma DNA plays an important role as biomarker. Cancers are genetically unstable and heterogeneous even within an individual's tumour. Furthermore, genetic instability enhances the efficiency of cancer development and the applied therapy might even induce defects in the cancer genome, so that it can lose their hitting target. Such a dynamic resistance to therapies has been shown for many tumour types with sometimes more than one resistance mechanism in the same patient [Beckman *et al.* (2012)]. However, it has been reported that not all patients with progressive metastatic disease seem to release enough tumour DNA into the circulation, so that quantities can be detected [Heitzer *et al*. (2013a)]. Such incidences make it important to develop a screening method that identifies patients where tumour-specific biomarkers can be most likely identified. The approach working as a pre-screening method has to be robust and would sort out the samples, which show low chromosomal alterations.

Different reports have described non-invasive screening methods in the last few years [Forshew *et al.* (2012); Dawson *et al.* (2013); Zong *et al.* (2012); Leary *et al.* (2012)]. Massively parallel sequencing of plasma DNA derived from the maternal blood is already available in the clinic for the detection of fetal aneuploidy. Fan *et al.* (2012) recently described a non-invasive prenatal diagnosis methodology measuring the entire fetal genome. Their approach enables exome screening of inherited and *de novo* genetic diseases.

Using the same approach, that is massively parallel sequencing of plasma DNA, identification of chromosomal alterations in the circulation of different cancer types including breast and

colorectal cancer has been described [Dawson *et al.* (2013); Heitzer *et al.* (2013b); Leary *et al.* (2012)].

Murtaza *et al.* (2013) reported a method for analysing of acquired cancer therapy resistance by plasma DNA sequencing. They performed exome sequencing from metastatic plasma samples spanning several courses of treatment to study the resistance to systemic treatment. The method uses commercially available kits to prepare the library and the HiSeq2500 sequencing platform. They were able to gain an overview of the mutation changes due to tracking genomic evolution in response to cancer treatment. However, whole exome sequencing enables a detailed analysis of the whole genome that is cost and time intensive. Thus, all methods mentioned above are not suitable for a fast and cheap screening test for numerous cancer samples.

On the other hand, FAST-SeqS offers the option of a fast screening method that gives an overview of chromosomal aberrations instead of a detailed genome analysis. It was originally published for the non-invasive detection of fetal chromosomal aneuploidy by massively parallel sequencing. Circulating fetal DNA derived from maternal plasma was used for detection of trisomy 21 (Down syndrome), trisomy 18 (Edwards syndrome) and trisomy 13 (Patau syndrome). FAST-SeqS implements two single PCRs with a specific primer pair to prepare libraries that are sequenced using NGS technologies. The new innovation that makes FAST-SeqS different from other library preparation systems is the use of specific primers, which anneal to a subset of repeated sequences throughout the genome. This technique increases on one hand the throughput and lowers also costs compared to whole-genome sequencing tests.

A defined number of fragments is amplified using LINE primers that target the repeated regions, which are more or less distributed randomly throughout the genome. This avoids the need for end-repair, terminal 3'-A addition, and ligation to adapters, which are necessary steps for common library preparation methods. Although FAST-SeqS cannot identify structural interchromosomal rearrangements such as translocations and determine the copy number changes of every distinct region of the genome, it can help to decide whether a sample should be analysed in more detail.

Kinde *et al.* (2012) reported the method for detection of fetal aneuploidies. However, the detection of chromosomal abnormalities within the genome of cancer patients was not reflected in their study. Therefore, the aim of this study was the implementation of this new method for analysis of chromosomal copy number variations of plasma DNA from cancer patients. It was not of interest to measure the location or extent of chromosomal aberrations or the clinical effects. Furthermore, this method captures only the chromosomal CNV at a certain time and does not provide any information about the tumour development. However, this screening method should offer a decision making support and give only an overview of chromosomal aberrations of cancer patients.

The study was conducted among men with prostate cancer, women with breast cancer and one male patient with colon cancer. Furthermore, genomic DNA derived from cancer cell lines were analysed in mixing experiments to test the sensitivity of the screening method. The protocol was adapted from the previously published FAST-SeqS from Kinde *et al.* (2012). Some slight modification had been made such as alter the number of PCR cycles for library preparation. The original two cycles in the first PCR were increased to five cycles and the formerly thirteen cycles for the second PCR were increased to fifteen cycles for cell line DNA and to eighteen cycles for plasma DNA. The higher number of PCR cycles were needed, because cell-free circulating DNA in the blood circulation have only a limited DNA amount.

The first PCR used one forward and one reverse primer, that both target LINE specific sequences and have a priming site attached for the sequencing primers used in the second PCR. The primer sequences for the second PCR include the priming side for the sequencing primer of the first primer pair and an adaptor sequence, which allowed the amplicons to attach to the flow cells of Illumina MiSeq system. The reverse primers contain in addition specific index sequences for the sequencing platform, which enable the identification of single samples from a pool of up to 24 samples after sequencing. The length of the priming side is of crucial importance, as a too short sequence resulted in no amplification of the PCR product of the first PCR. A total length of 34 bp for the priming side appeared to be sufficient to guarantee an efficient amplification of the first PCR product.

The prepared libraries were sequenced with the Illumina MiSeq system, in contrast to the study of Kinde *et al.* (2012), who used a HiSeq system that enables a much higher throughput. Therefore, the sequenced read counts per sample were lower in comparison to their study; 1,343,382 to 16,015,347 reads per sample compared to this study with 925,843 to 7,267 reads per sample. Although, the numbers of reads were lower, reliable and meaningful results were obtained, which can attract smaller laboratories to establish this new method for cancer pre-screening as it has dramatically reduced run times compared to the Illumina HiSeq.

Invasive sample collection from tumour tissue was not necessary, as plasma DNA was used for the analysis. Depending on the DNA amount used in the first PCR, different concentrations of libraries were measured after the two amplifications by PCR. DNA of 28 patients with breast cancer, 27 patients with prostate cancer and one patient with colon cancer were analysed by the adapted FAST-SeqS method. The chromosomal copy number status was determined by calculating the z-scores for single chromosome arms and the entire genome. Cancer samples showed several gains and losses in different chromosome arms as expected compared to the control samples.

The metastatic samples showed in general much more gains and losses of chromosome arms than non-metastatic samples. Furthermore, prostate cancer patient showed frequently aberrations of the X-chromosome. The X-chromosome is known to include the androgen

receptor, which plays an important role in the resistance of androgen deprivation therapy. Therefore, mutations on the X-chromosome are relatively common among prostate cancer patients [Gelman (2002)]. The prostate cancer patients as well as breast cancer and colon cancer patients also showed frequent copy number changes on the chromosome 8, which is known to be involved in cancer.

The previously published PlasmaSeq method by Heitzer *et al.* (2013b) was applied earlier on the prostate cancer and colon cancer samples for another study (unpublished results to this day). Those results were compared to confirm the copy number aberrations observed with FAST-SeqS. The breast cancer results were confirmed with CNV profiles obtained with array CGH. The CNV profiles of PlasmaSeq as well as array CGH showed an agreement with the results of FAST-SeqS. Yet, some discrepancies occurred, which can be caused by the simultaneous event of gains and losses in the same chromosome arm leading to false-negative z-scores. Other CNV were too small compared to the rest of the chromosome arm, that they were not detectable.

Heitzer *et al.* (2013b) published recently a new sequencing approach named PlasmaSeq employing a benchtop high-throughput sequencing instrument. Using this instrument whole genome sequencing from plasma DNA was performed to generate genome-wide copy number profiles of cancer patients with low cost and time effort. Shotgun libraries were prepared following the manufacturer's instructions using the TruSeq DNA LT Sample preparation Kit (Illumina, USA) for the PlasmaSeq libraries. However, some slight modifications were made such as lower amount of input DNA, skipping the fragmentation step, and PCR amplification with 25 cycles. Using the same sequencing platform and sequencing conditions, the main difference between PlasmaSeq and FAST-SeqS is the library preparation step. While FAST-SeqS only amplifies a discrete subset of LINE sequences throughout the entire genome, PlasmaSeq enables a whole genome analysis from plasma DNA, which is more time and cost intensive.

PlasmaSeq showed a higher sensitivity for the detection of small CNVs as provided in figure 49. With this sensitive technique, small gains and losses such as in chromosome 3 and 6 were identified that could not be detected with FAST-SeqS.

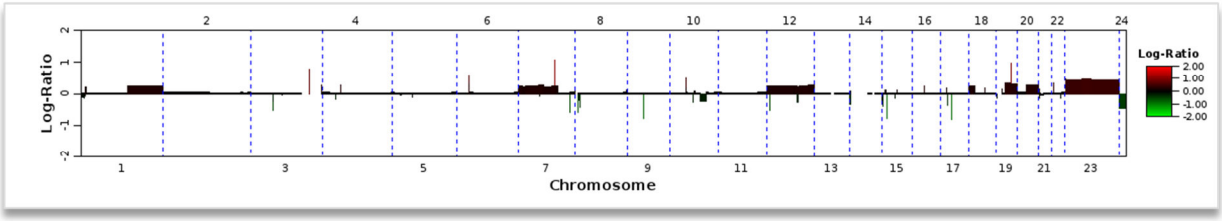


Figure 49: Log2-ratio results of colon cancer sample c38 from PlasmaSeq sequencing.

FAST-SeqS was also not able to detect the small CNVs that were spotted in the breast cancer samples by array CGH due to its much lower sensitivity compared to array CGH. However, the results of FAST-SeqS were comparable to outcomes of PlasmaSeq and array CGH.

Several copy number changes in circulating cell free DNA derived from cancer patients, which were already identified with array CHG or PlasmaSeq, were confirmed using the genome-wide z-scores, which was calculated based on the z-score of all chromosome arms.

Cancer patient usually show higher concentrations of cell free DNA, because the plasma DNA is derived from normal, non-tumour cells as well as from primary tumour or metastases. Highly sensitive methods are needed to detect the tumour DNA among the non-tumour DNA in the plasma. Therefore, tests for sensitivity were conducted using four cancer cell lines. Mixing experiments with DNA from normal individuals and from cancer cell lines should evaluate the sensitivity and limits of FAST-SeqS. Using the read counts obtained from the sequencing run, chromosomal copy number status was determined by calculating the z-scores for chromosomes, single chromosome arms and the entire genome.

The dilution series showed that tumour DNA concentrations up to a ratio of 15% are sufficient for a meaningful analysis. The diluted samples with DNA quantities less than 15% of cancer cell line DNA showed a few gains and losses, but the quantity deviates strongly from the actual number of aberrations. Especially the fragmented HCT116 and HepG2 cell line showed for the dilution-samples with DNA quantities less than 20% randomly gains and losses, so that the results cannot be completely reliable as the z-score should steadily decline with the decrease of tumour DNA. Many factors could have contributed to the results: unaccuracy while pitpetting small amounts of tumour DNA and measuring the DNA end concentration for sequencing with the Bioanalyzer. Furthermore, at the beginning of the study all non-fragmented DNA samples for the dilution series with a sequencing index higher than 14 were amplified in the second PCR with shorter primers resulting in a lower concentration of the PCR product. In general, the z-scores based on the chromosome arms showed a better dilution series than the results based on the entire chromosome. Furthermore, dilution experiments suggested that the genome-wide z-score detects aneuploidy even at tumour DNA concentrations of only 1% for some cell lines, whereas for other cell lines a tumour DNA concentration of 50% is needed to detect CNV based on a genome-wide z-score.

Kinde *et al.* (2012) described FAST-SeqS as a new screening method for fetal aneuploidy with the advantages in ease of implementation, cost, analysis time, and throughput. In this study their outcome was confirmed and it has been shown that the method could also be used for screening tests for chromosomal aberrations in cancer patients.

Especially attractive features of this new method include the speed (library preparation, approximately 4 hours; sequencing of 150 bp single reads, approximately 12 hours; calculation of z-scores, approximately 30 minutes) and the reduced costs with which the

CNV screening by FAST-SeqS can be performed and are the main advantages to implement the method into tumour screening. A small amount of DNA (less than 50 ng) is enough for sufficient sequencing reads that can be interpreted.

The disadvantage of FAST-SeqS is the reduced resolution, so that no mutations or structural intrachromosomal rearrangements can be identified. Simultaneous gains and losses in one chromosome arm cannot be detected and leads to an inconspicuous z-score. Furthermore, the origin of the plasma DNA is not clear, so that observed changes in the plasma can be related to the primary tumour or to any of the metastatic sites. In fact, it is currently unknown whether all tumour cells release DNA into the plasma equally and by which factors it is influenced.

In summary, the results of this study suggest that chromosomal information can be derived from FAST-SeqS, which may facilitate the pre-screening of cancer patients. Chromosomal copy number variations of tumour samples can be detected with the new screening method that analyses only a defined number of LINE sequences. Its advantage over whole genome sequencing is caused by the simple implementation, lower cost, decreased analysis time, and higher throughput. Further studies involving a large number of cancer patients will reveal whether it will perform as well as whole genome sequencing in the clinic. As results have been obtained for colon, breast, and prostate cancer patients, this method may also be applicable to other tumour types.

The importance of plasma DNA samples for diagnostics and its analyses is steadily increasing. However, the entire genome of all cancer patients cannot be analysed in detail due to cost and time limitations. Whether FAST-SeqS will be implemented in the management of cancer patients has to be further evaluated. Though, it could potentially aid in identifying patients more or less likely showing chromosomal aberrations and present a first overview of cancer related chromosomal gains and losses. The simplicity, the costs, and the reduced time of FAST-SeqS are attractive and might ease the clinical translation.

REFERENCES

Angelis, R. de; Sant, M.; Coleman, M. P.; Francisci, S.; Baili, P.; Pierannunzio, D. *et al.* (2013): Cancer survival in Europe 1999-2007 by country and age: results of EUROCARE-5-a population-based study. In *Lancet Oncology 2013. DOI:* 10.1016/S1470-2045(13)70546-1.

Anker, P.; Stroun, M.; Maurice, P. A. (1976): Spontaneous extracellular synthesis of DNA released by human blood lymphocytes. In *Cancer Res* 36 (8), pp. 2832–2839.

Arcila, Maria; Lau, Christopher; Nafa, Khedoudja; Ladanyi, Marc (2011): Detection of KRAS and BRAF Mutations in Colorectal Carcinoma. In *The Journal of Molecular Diagnostics* 13 (1), pp. 64–73. DOI: 10.1016/j.jmoldx.2010.11.005.

Beckman, R. A.; Schemmann, G. S.; Yeang, C.-H. (2012): Impact of genetic dynamics and single-cell heterogeneity on development of nonstandard personalized medicine strategies for cancer. In *Proceedings of the National Academy of Sciences* 109 (36), pp. 14586–14591. DOI: 10.1073/pnas.1203559109.

Beckman Coulter (2009): Agencourt AMPure XP PCR Purification. 000387v001.

BioMed Central (2005): Somatic mutation and gain of copy number of PIK3CA in human breast cancer. With assistance of Guojun Wu, Mingzhao Xing, Elizabeth Mambo, Xin Huang, Junwei Liu, Zhongmin Guo et al. Available online at http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1242128/#!po=3.84615, updated on 1/1/2005, checked on 11/6/2013.

BIO-RAD (2014): PCR Primer and Probe Chemistries, checked on 1/27/2014.

Blackburn, E. H. (2005): Telomerase and Cancer: Kirk A. Landon - AACR Prize for Basic Cancer Research Lecture. In *Molecular Cancer Research* 3 (9), pp. 477–482. DOI: 10.1158/1541-7786.MCR-05-0147.

Bozic, I.; Antal, T.; Ohtsuki, H.; Carter, H.; Kim, D.; Chen, S. *et al.* (2010): Accumulation of driver and passenger mutations during tumor progression. In *Proceedings of the National Academy of Sciences* 107 (43), pp. 18545–18550. DOI: 10.1073/pnas.1010978107.

Brown, T. A. (2010): Gene cloning and DNA analysis. An introduction. 6th ed. Oxford: Wiley-Blackwell.

Chan, K. C. A.; Jiang, P.; Zheng, Y. W. L.; Liao, G. J. W.; Sun, H.; Wong, J. *et al.* (2013): Cancer Genome Scanning in Plasma: Detection of Tumor-Associated Copy Number Aberrations, Single-Nucleotide Variants, and Tumoral Heterogeneity by Massively Parallel Sequencing. In *Clinical Chemistry* 59 (1), pp. 211–224. DOI: 10.1373/clinchem.2012.196014.

Chang, Christine; Chia, Rhu-Hsin; Wu, Tsu-Lan; Tsao, Kuo-Chien; Sun, Chien-Feng; Wu, James (2003): Elevated cell-free serum DNA detected in patients with myocardial infarction. In *Clin Chim Acta* 327 (1-2), pp. 95–101.

Costantini, S.; Di Bernardo, G.; Cammarota, M.; Castello, G.; Colonna, G. (2013): Gene expression signature of human HepG2 cell line. In *Gene* 518 (2), pp. 335–345. DOI: 10.1016/j.gene.2012.12.106.

Crowley-Weber, C. L. (2002): Development and molecular characterization of HCT-116 cell lines resistant to the tumor promoter and multiple stress-inducer, deoxycholate. In *Carcinogenesis* 23 (12), pp. 2063–2080. DOI: 10.1093/carcin/23.12.2063.

Dawson, Sarah-Jane; Tsui, Dana W.Y.; Murtaza, Muhammed; Biggs, Heather; Rueda, Oscar M.; Chin, Suet-Feung *et al.* (2013): Analysis of Circulating Tumor DNA to Monitor Metastatic Breast Cancer. In *N Engl J Med* 368 (13), pp. 1199–1209. DOI: 10.1056/NEJMoa1213261.

Einstein, Michael (2012): The battle for sequencing supremacy. In *Nature Biotechnology* 30, pp. 1023–1026. DOI: 10.1038/nbt.2412.

Elgar, Greg; Vavouri, Tanya (2008): Tuning in to the signals: noncoding sequence conservation in vertebrate genomes. In *Trends in Genetics* 24 (7), pp. 344–352. DOI: 10.1016/j.tig.2008.04.005.

Fan, Christina; Gu, Wei; Wang, Jianbin; Blumenfeld, Yair; El-Sayed, Yasser; Quake, Stephen (2012): Non-invasive prenatal measurement of the fetal genome. In *Nature Biotechnology* 487 (7407), pp. 320–324. DOI: 10.1038/nature11251.

Forshew, T.; Murtaza, M.; Parkinson, C.; Gale, D.; Tsui, D. W. Y.; Kaper, F. *et al.* (2012): Noninvasive Identification and Monitoring of Cancer Mutations by Targeted Deep Sequencing of Plasma DNA. In *Science Translational Medicine* 4 (136), pp. 136ra68. DOI: 10.1126/scitranslmed.3003726.

Gelman, Edward (2002): Molecular Biology of the Androgen Receptor. In *journal of Clinical Oncology* 20 (13), pp. 3001–3015.

Gerlinger, Marco; Rowan, Andrew J.; Horswell, Stuart; Larkin, James; Endesfelder, David; Gronroos, Eva *et al.* (2012): Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. In *N Engl J Med* 366 (10), pp. 883–892. DOI: 10.1056/NEJMoa1113205.

Ghorbian, Saeid; Ardekani, Ali (2012): Non-Invasive Detection of Esophageal Cancer using Genetic Changes in Circulating Cell-Free DNA. In *Avicenna J Med Biotechnol.* 4 (1), pp. 3–13.

Hanahan, Douglas; Weinber, Robert (2011): Hallmarks of Cancer: The Next Generation. In *Cell* 144 (5), pp. 646–674. DOI: 10.1016/j.cell.2011.02.013.

Hancks, Dustin C.; Kazazian, Haig H. (2012): Active human retrotransposons: variation and disease. In *Current Opinion in Genetics & Development* 22 (3), pp. 191–203. DOI: 10.1016/j.gde.2012.02.006.

Heitzer, Ellen; Auer, Martina; Hoffmann, Eva Maria; Pichler, Martin; Gasch, Christin; Ulz, Peter *et al.* (2013a): Establishment of tumor-specific copy number alterations from plasma DNA of patients with cancer. In *Int. J. Cancer* 133 (2), pp. 346–356. DOI: 10.1002/ijc.28030.

Heitzer, Ellen; Ulz, Peter; Belic, Jelena; Gutschi, Stefan; Quehenberger, Franz; Fischereder, Katja *et al.* (2013b): Tumor-associated copy number changes in the circulation of patients with prostate cancer identified through whole-genome sequencing. In *Genome Med* 5 (4), p. 30. DOI: 10.1186/gm434.

Holland, Andrew J.; Cleveland, Don W. (2012): Losing balance: the origin and impact of aneuploidy in cancer. In *EMBO Rep* 13 (6), pp. 501–514. DOI: 10.1038/embor.2012.55.

Illumina (2010): Illumina Sequencing Technology. Illumina, updated on 10/11/2010, checked on 1/8/2014.

Illumina (2013): TruSeq DNA Sample Preparation Kits. Available online at http://www.illumina.com/products/truseq_dna_sample_prep_kits.ilmn, checked on 11/5/2013.

Jurka, Jerzy; Kapitonov, Vladimir V.; Kohany, Oleksiy; Jurka, Michael V. (2007): Repetitive Sequences in Complex Genomes: Structure and Evolution. In *Annu. Rev. Genom. Human Genet.* 8 (1), pp. 241–259. DOI: 10.1146/annurev.genom.8.080706.092416.

Kinde, Isaac; Papadopoulos, Nickolas; Kinzler, Kenneth W.; Vogelstein, Bert; Veitia, Reiner Albert (2012): FAST-SeqS: A Simple and Efficient Method for the Detection of Aneuploidy by Massively Parallel Sequencing. In *PLoS ONE* 7 (7), pp. e41162. DOI: 10.1371/journal.pone.0041162.

Klug, William S.; Cummings, Michael R. (2003): Concepts of genetics. 7th ed. Upper Saddle River, N.J: Prentice Hall.

Kumar, S. (2006): Circulating Cell-Free DNA in Plasma/Serum of Lung Cancer Patients as a Potential Screening and Prognostic Tool. In *Clinical Chemistry. DOI:* 10.1373/clinchem.2005.062893.

Lam, Hugo; Clark, Michael; Chen, Rui; Chen, Rong; Natsoulis, Georges; O'Huallachain, Maeve *et al.* (2012): Performance comparison of whole-genome sequencing platforms. In *Nature Biotechnology* 30, pp. 78–82. DOI: 10.1038/nbt.2065.

Leary, R. J.; Sausen, M.; Kinde, I.; Papadopoulos, N.; Carpten, J. D.; Craig, D. *et al.* (2012): Detection of Chromosomal Alterations in the Circulation of Cancer Patients with Whole-Genome Sequencing. In *Science Translational Medicine* 4 (162), pp. 162ra154. DOI: 10.1126/scitranslmed.3004742.

Levenson, Anait S.; Jordan, V. Craig (1997): MCF-7: The First Hormone-responsive Breast Cancer Cell Line. In *Cancer Res* 57, pp. 3071–3078. Available online at http://cancerres.aacrjournals.org/content/57/15/3071.full.pdf.

Li, H.; Durbin, R. (2009): Fast and accurate short read alignment with Burrows-Wheeler transform. In *Bioinformatics* 25 (14), pp. 1754–1760. DOI: 10.1093/bioinformatics/btp324.

Liu, Lin; Li, Yinhu; Li, Siliang; Hu, Ni; He, Yimin; Pong, Ray *et al.* (2012): Comparison of Next-Generation Sequencing Systems. In *Journal of Biomedicine and Biotechnology* 2012 (7), pp. 1–11. DOI: 10.1155/2012/251364.

Lo, Y. M. D.; Tein, M.; Lau, T.; Haines, C.; Leung, T.; Poon, P. *et al.* (1998): Quantitative Analysis of Fetal DNA in Maternal Plasma and Serum: Implications for Noninvasive Prenatal Diagnosis. In *The American Journal of Human Genetics* 62 (4), pp. 768–775. DOI: 10.1086/301800.

Loman, Nicholas; Misra, Raju; Dallman, Timothy; Constantinidou, Chrystala; Gharbia, Saheer; Wain, John; Pallen, Mark (2012): Performance comparison of benchtop high-throughput sequencing platforms. In *Nature Biotechnology* 30, pp. 434–439. DOI: 10.1038/nbt.2198.

McClintock, B. (1984): The significance of responses of the genome to challenge. In *Science* 226 (4676), pp. 792–801. DOI: 10.1126/science.15739260.

McGranahan, Nicholas; Burrell, Rebecca A.; Endesfelder, David; Novelli, Marco R.; Swanton, Charles (2012): Cancer chromosomal instability: therapeutic and diagnostic challenges. In *EMBO Rep* 13 (6), pp. 528–538. DOI: 10.1038/embor.2012.61.

Murtaza, Muhammed; Dawson, Sarah-Jane; Tsui, Dana W. Y.; Gale, Davina; Forshew, Tim; Piskorz, Anna M. *et al.* (2013): Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. In *Nature* 497 (7447), pp. 108–112. DOI: 10.1038/nature12065.

Niforou, K. M.; Anagnostopoulos, A. K.; Vougas, K.; Kittas, C.; Gorgoulis, V. G.; Tsangaris, G. T. (2008): The proteome profile of the human osteosarcoma U2OS cell line. In *Cancer Genomics and Proteomics.* 5 (1), pp. 63–78.

O'Huallachain, M.; Karczewski, K. J.; Weissman, S. M.; Urban, A. E.; Snyder, M. P. (2012): Extensive genetic variation in somatic human tissues. In *Proceedings of the National Academy of Sciences* 109 (44), pp. 18018–18023. DOI: 10.1073/pnas.1213736109.

Olivier, M.; Hollstein, M.; Hainaut, P. (2010): TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use. In *Cold Spring Harbor Perspectives in Biology* 2 (1), pp. a001008. DOI: 10.1101/cshperspect.a001008.

Pontén, J.; Saksela, E. (1967): Two established in vitro cell lines from human mesenchymal tumours. In *Int J Cancer* 2 (5), pp. 434–447.

Rainer, T. H. (2003): Prognostic Use of Circulating Plasma Nucleic Acid Concentrations in Patients with Acute Stroke. In *Clinical Chemistry* 49 (4), pp. 562–569. DOI: 10.1373/49.4.562.

Ricke, R. M.; van Deursen, J. M. (2013): Aneuploidy in health, disease, and aging. In *The Journal of Cell Biology* 201 (1), pp. 11–21. DOI: 10.1083/jcb.201301061.

Rodić, Nemanja; Burns, Kathleen H.; Rosenberg, Susan M. (2013): Long Interspersed Element–1 (LINE-1): Passenger or Driver in Human Neoplasms? In *PLoS Genet* 9 (3), pp. e1003402. DOI: 10.1371/journal.pgen.1003402.

Roy-Engel, Astrid, M. (2012): LINEs, SINEs and other retroelements: do birds of a feather flock together? In *Front Biosci* 17 (1), p. 1345. DOI: 10.2741/3991.

Sanger, F.; Nicklen, S.; Coulson, A. R. (1977): DNA sequencing with chain-terminating inhibitors. In *Proceedings of the National Academy of Sciences* 74 (12), pp. 5463–5467.

Soule, H. D.; Vazguez, J.; Long, A.; Albert, S.; Brennan, M. (1973): A human cell line from a pleural effusion derived from a breast carcinoma. In *J Natl Cancer Inst* 51 (5), pp. 1409–1416.

Statistics Austria (2013): Cancer incidence. Statistics Austria. Available online at http://www.statistik.at/web_en/statistics/health/cancer_incidence/index.html, updated on 12/3/2013, checked on 1/21/2014.

Swarup, Vishnu; Rajeswari, M. R. (2007): Circulating (cell-free) nucleic acids – A promising, non-invasive tool for early detection of several human diseases. In *FEBS Letters* 581 (5), pp. 795–799. DOI: 10.1016/j.febslet.2007.01.051.

The SAM/BAM Format Specification Working Group (2013): Sequence Alignment/ Map Format Specification. Available online at http://samtools.sourceforge.net/SAMv1.pdf, checked on 1/19/2014.

Vogelstein, B.; Papadopoulos, N.; Velculescu, V. E.; Zhou, S.; Diaz, L. A.; Kinzler, K. W. (2013): Cancer Genome Landscapes. In *Science* 339 (6127), pp. 1546–1558. DOI: 10.1126/science.1235122.

Yeung, T. M.; Gandhi, S. C.; Wilding, J. L.; Muschel, R.; Bodmer, W. F. (2010): Cancer stem cells from colorectal cancer-derived cell lines. In *Proceedings of the National Academy of Sciences* 107 (8), pp. 3722–3727. DOI: 10.1073/pnas.0915135107.

Zong, C.; Lu, S.; Chapman, A. R.; Xie, X. S. (2012): Genome-Wide Detection of Single-Nucleotide and Copy-Number Variations of a Single Human Cell. In *Science* 338 (6114), pp. 1622–1626. DOI: 10.1126/science.1229164.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

µl	microliter
1×TE buffer	10 mMTris and 1mM EDTA
bp	base pair
BWA	Burrows-Wheeler Alignment Tool
cfDNA	cell free DNA
CGH	comparative genomic hybridization
Chr	chromosome
CIN	chromosomal instability
CNV	copy number variation
CV	coefficient of variation
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
dsDNA	double stranded deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
g	gravitational acceleration
HT1	hybridization buffer
LINE	long interspersed elements
LTR	long terminal repeat
min	minutes
mM	millimolar
mRNA	messenger ribonucleic acid
n.f.	nuclease free
NaOH	sodium hydroxide
ng	nanogram
NGS	next generation sequencing

nM	nanomolar
ORF	open reading frame
PCR	polymerase chain reaction
pH	power of hydrogen or potential hydrogen
pM	picomolar
PSA	prostate specific antigen
RNA	ribonucleic acid
RT	room temperature
RT-PCR	real time polymerase chain reaction
SAM	sequence alignment /map
SINE	short interspersed elements
ssDNA	single stranded deoxyribonucleic acid
TE	transposable elements
Tris-Cl	Trishydroxymethylamino methane
UTR	untranslated region
WGA	whole- genome amplification

APPENDICES

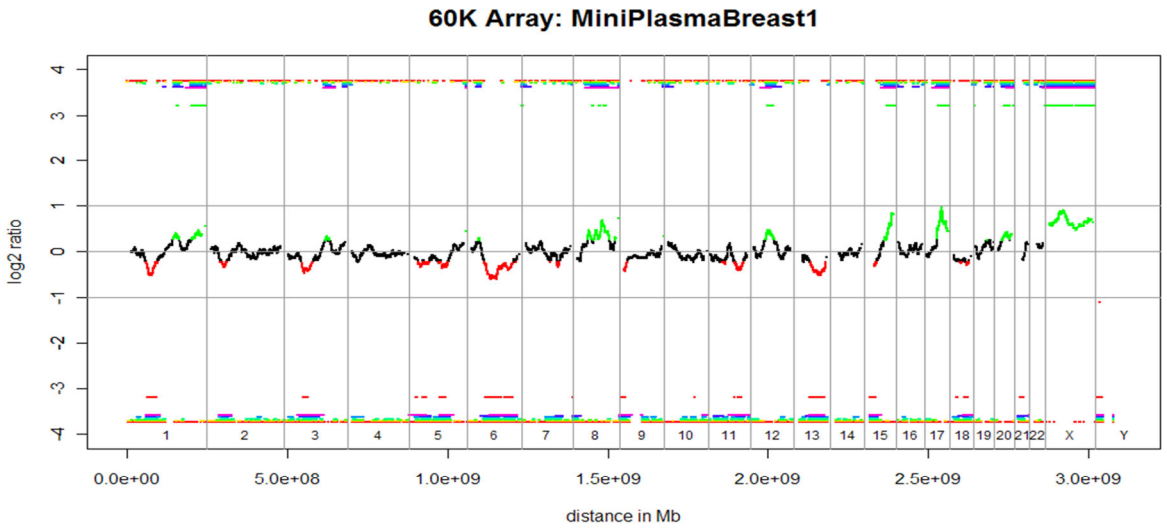
PRIMER SEQUENCES OF PCR1:

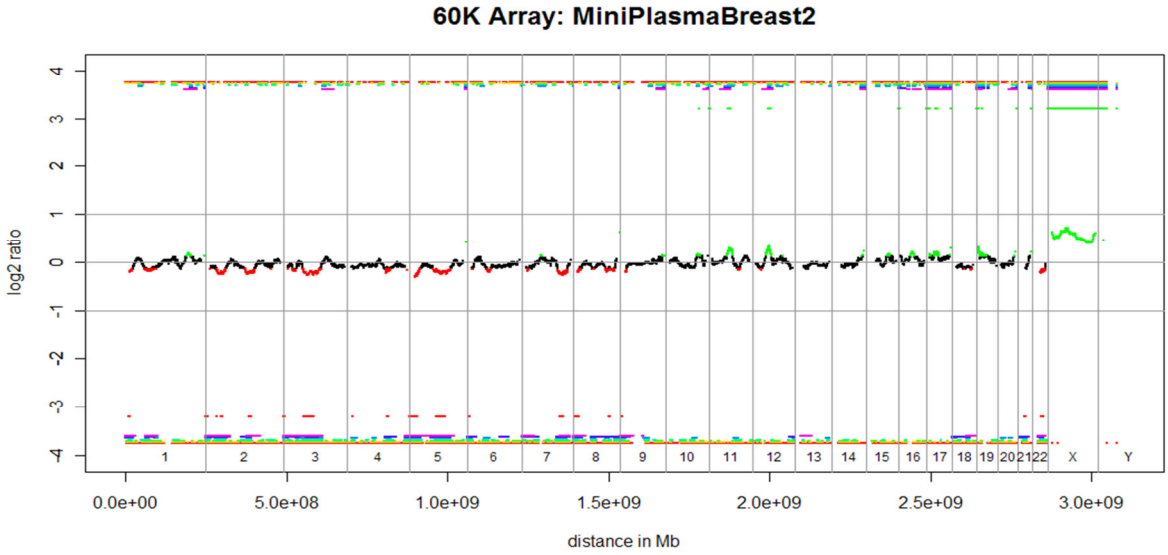
Primer	
Forward	TCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNNNNNNNNNACACAGGGAGGGGAACAT
Reverse	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTGCCATGGTGGTTTGCT

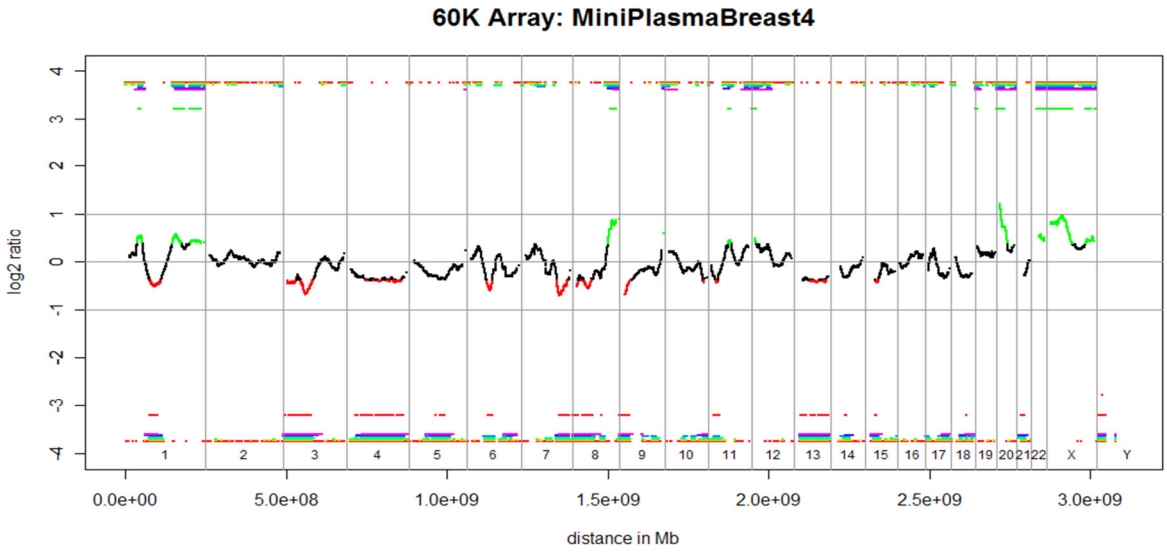
PRIMER SEQUENCES OF PCR 2:

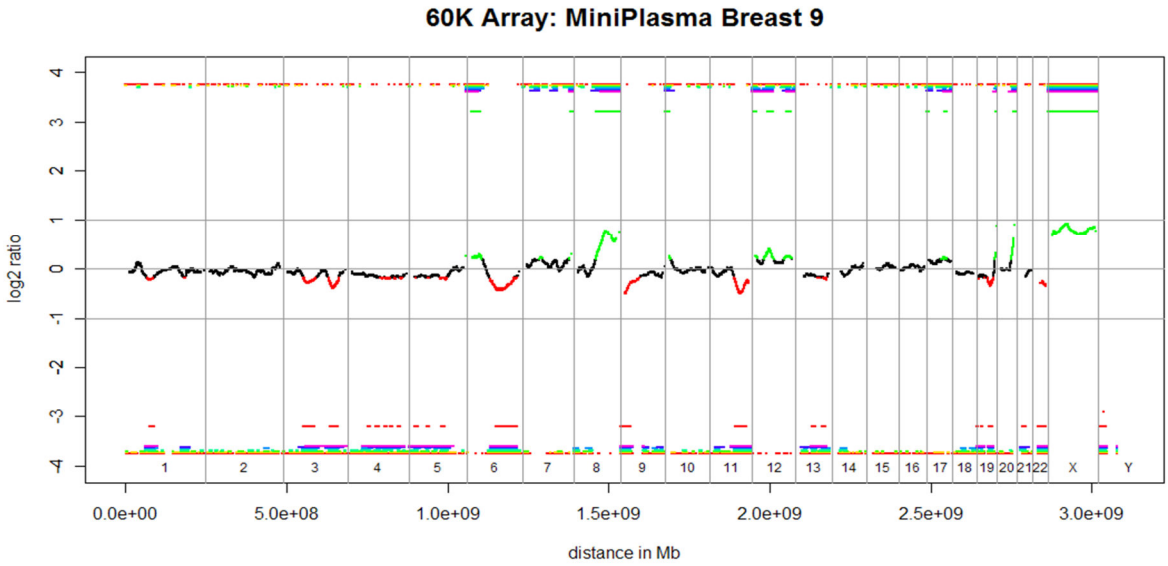
Primer	
Forward	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
Rev_Index1	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index2	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index3	CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index4	CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index5	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index6	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index7	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index8	CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index9	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index10	CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index11	CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index12	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index13	CAAGCAGAAGACGGCATACGAGATTTGACTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index14	CAAGCAGAAGACGGCATACGAGATGGAACTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index15	CAAGCAGAAGACGGCATACGAGATTGACATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index16	CAAGCAGAAGACGGCATACGAGATGGACGGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index18	CAAGCAGAAGACGGCATACGAGATGCGGACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index19	CAAGCAGAAGACGGCATACGAGATTTTCACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index20	CAAGCAGAAGACGGCATACGAGATGGCCACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index21	CAAGCAGAAGACGGCATACGAGATCGAAACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index22	CAAGCAGAAGACGGCATACGAGATCGTACGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index23	CAAGCAGAAGACGGCATACGAGATCCACTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index25	CAAGCAGAAGACGGCATACGAGATATCAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Rev_Index27	CAAGCAGAAGACGGCATACGAGATAGGAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

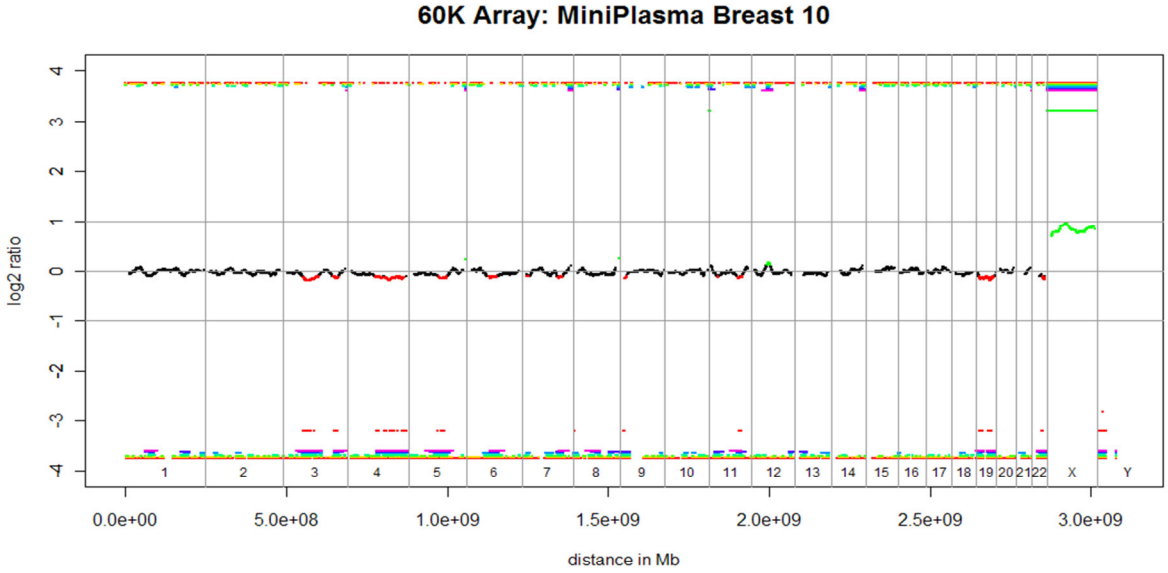
BREAST CANCER SAMPLES: ARRAY CGH PROFILES

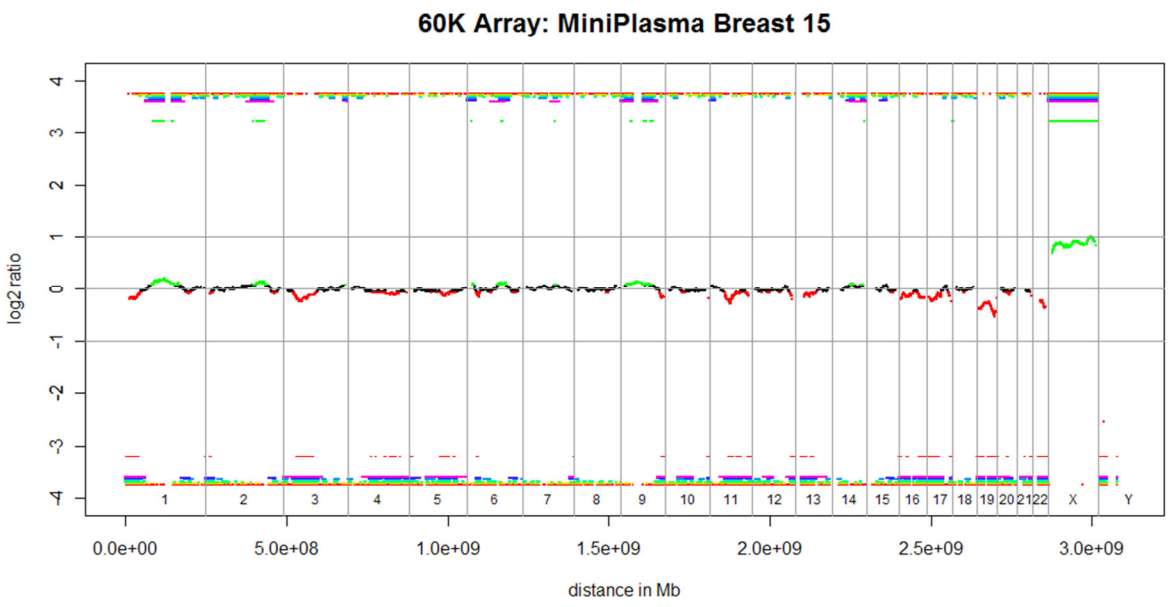


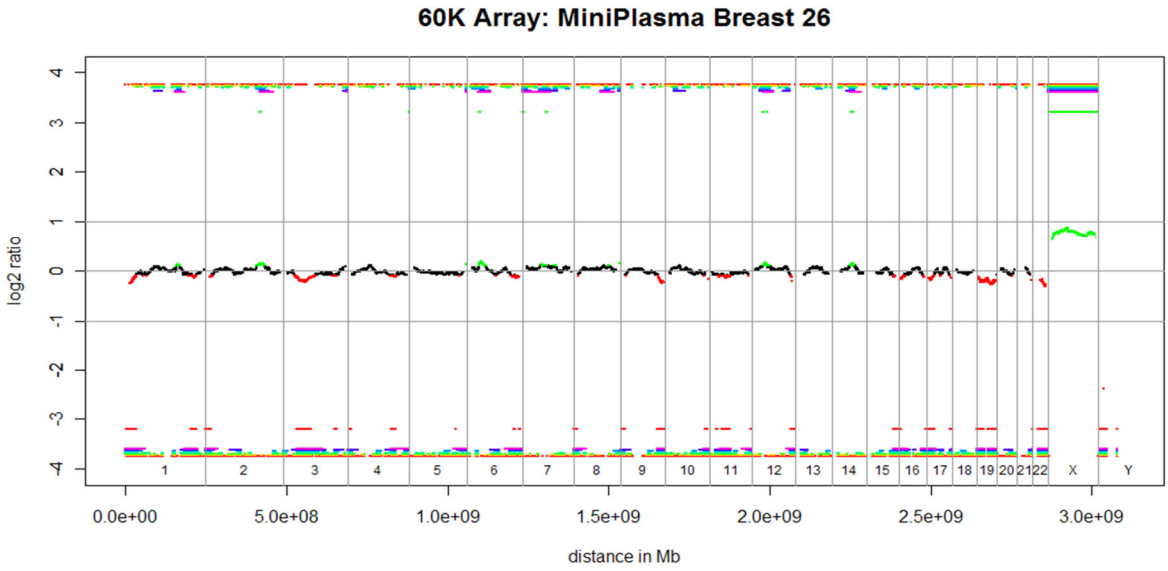


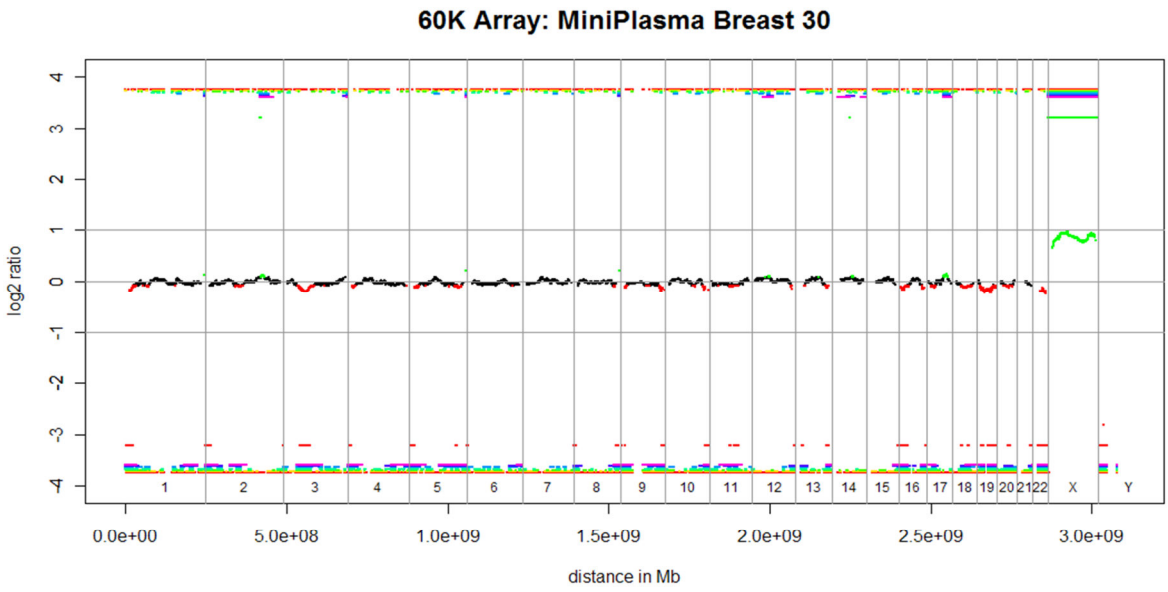


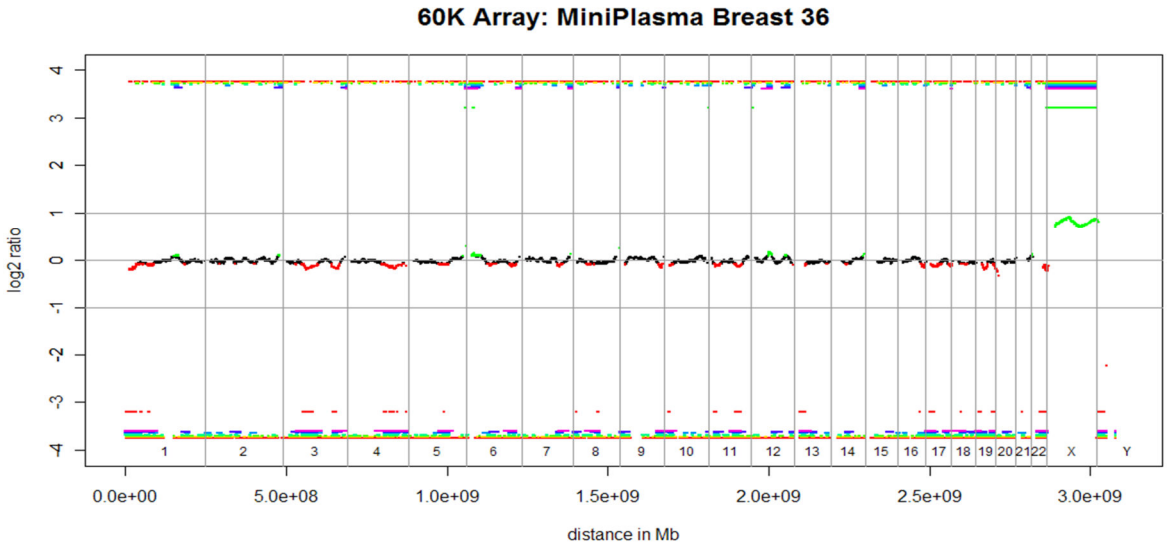

60K Array: MiniPlasma Breast 9
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

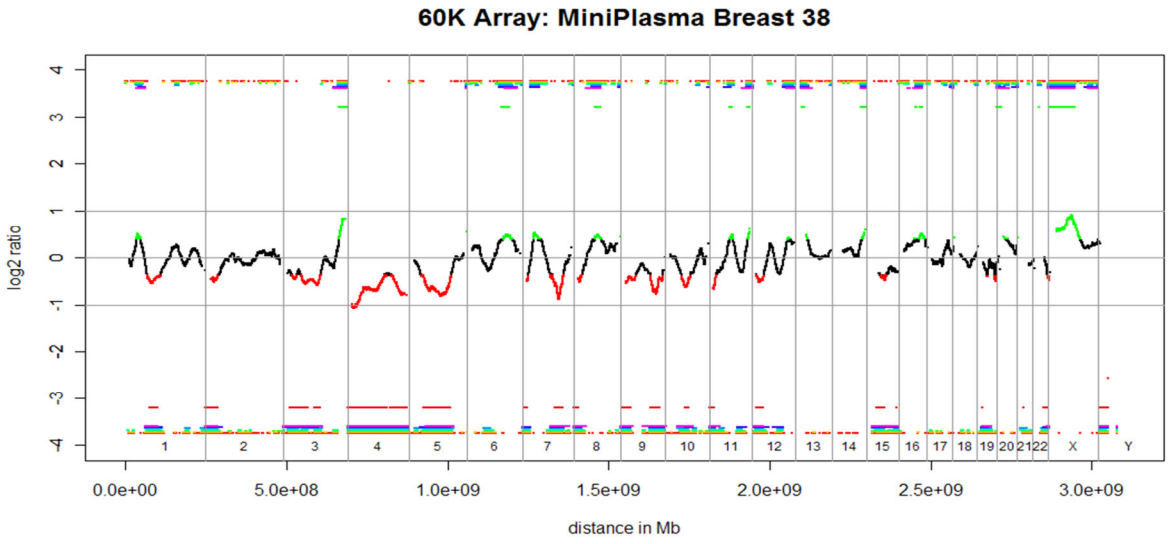

60K Array: MiniPlasma Breast 10
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

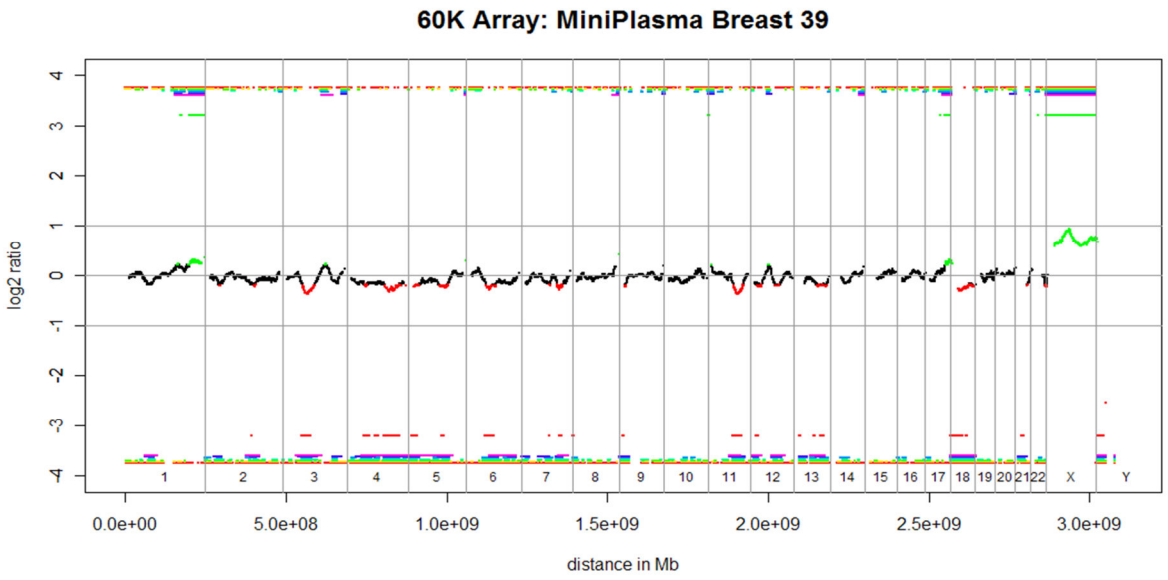

60K Array: MiniPlasma Breast 15
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

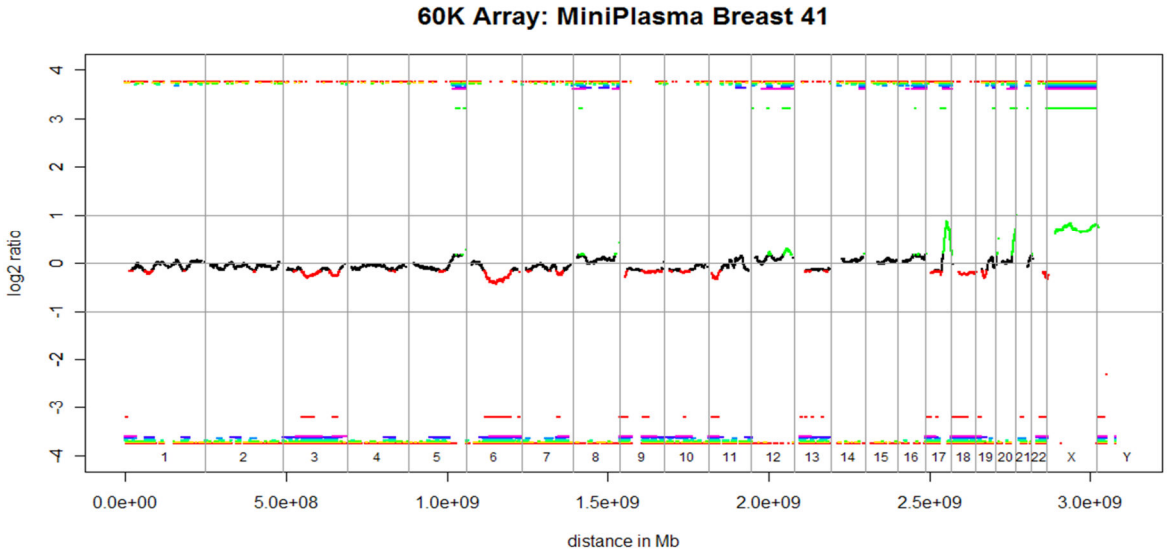

60K Array: MiniPlasma Breast 26
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

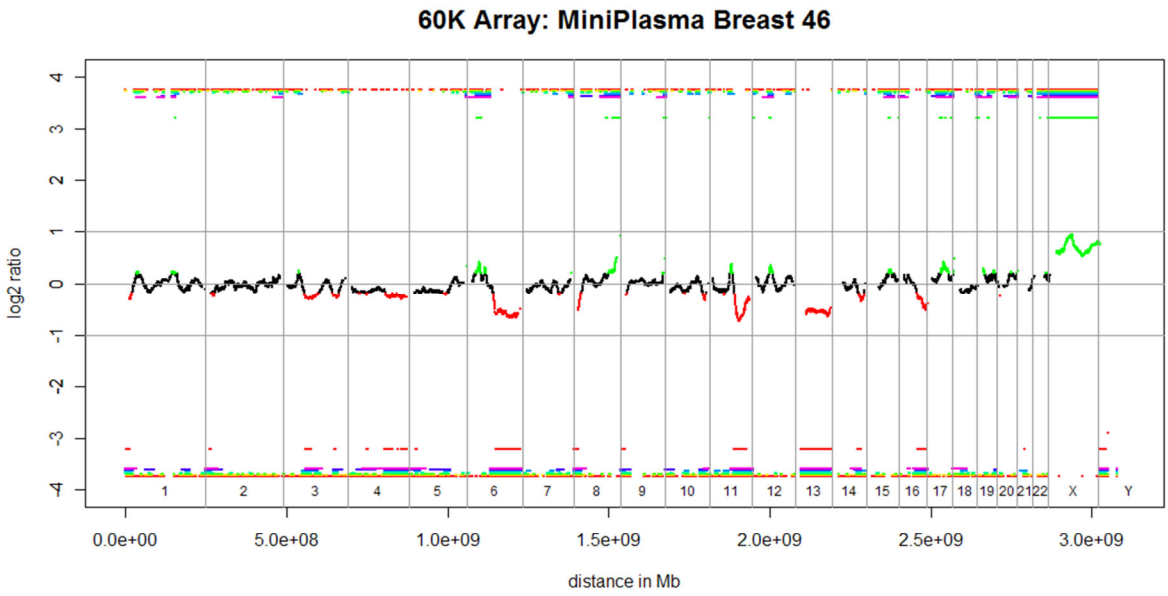

60K Array: MiniPlasma Breast 30
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

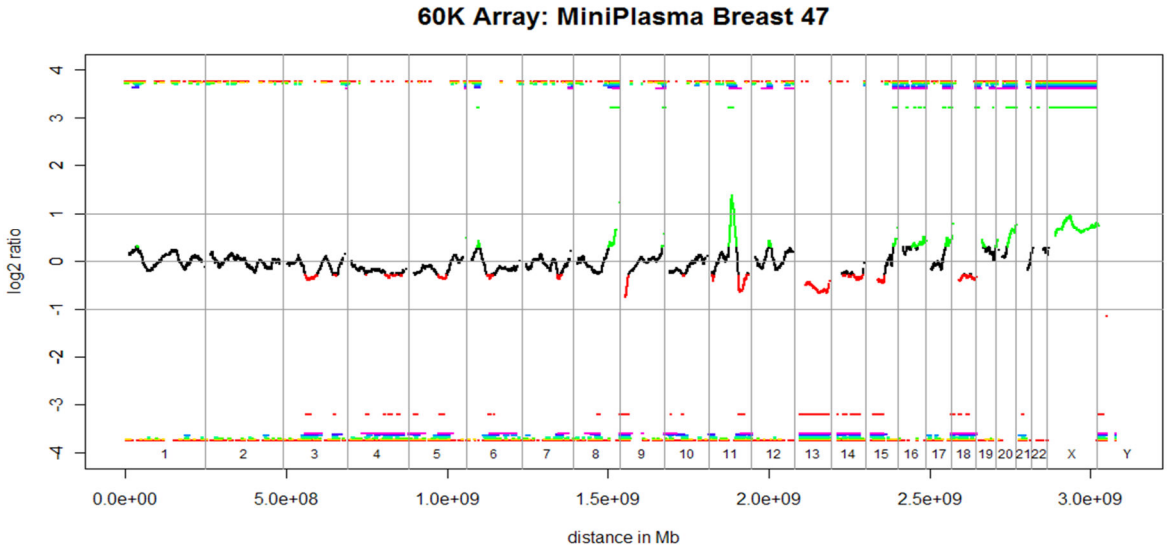

60K Array: MiniPlasma Breast 36
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

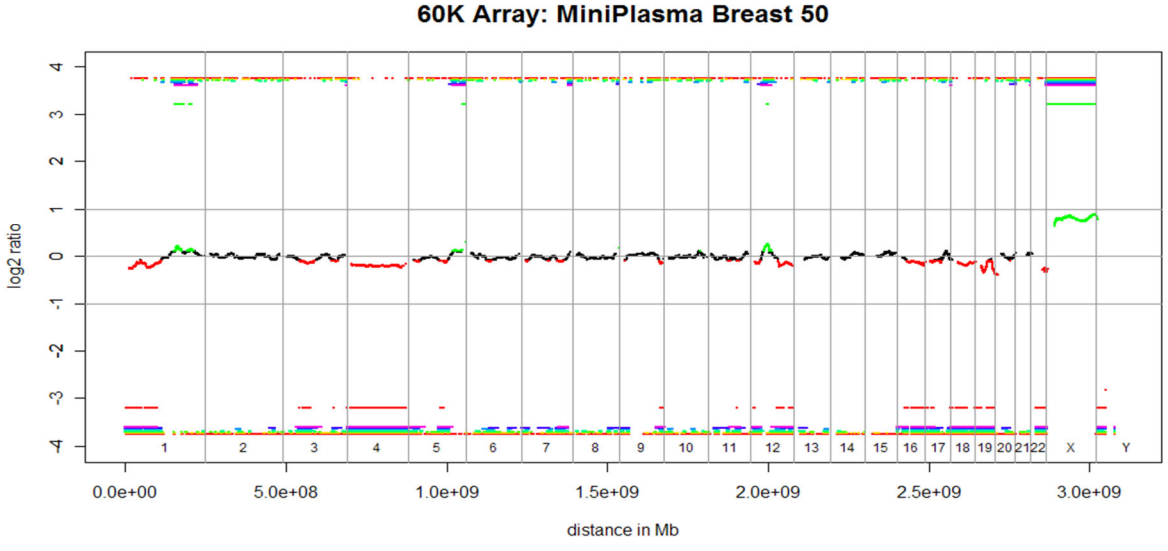

60K Array: MiniPlasma Breast 38
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

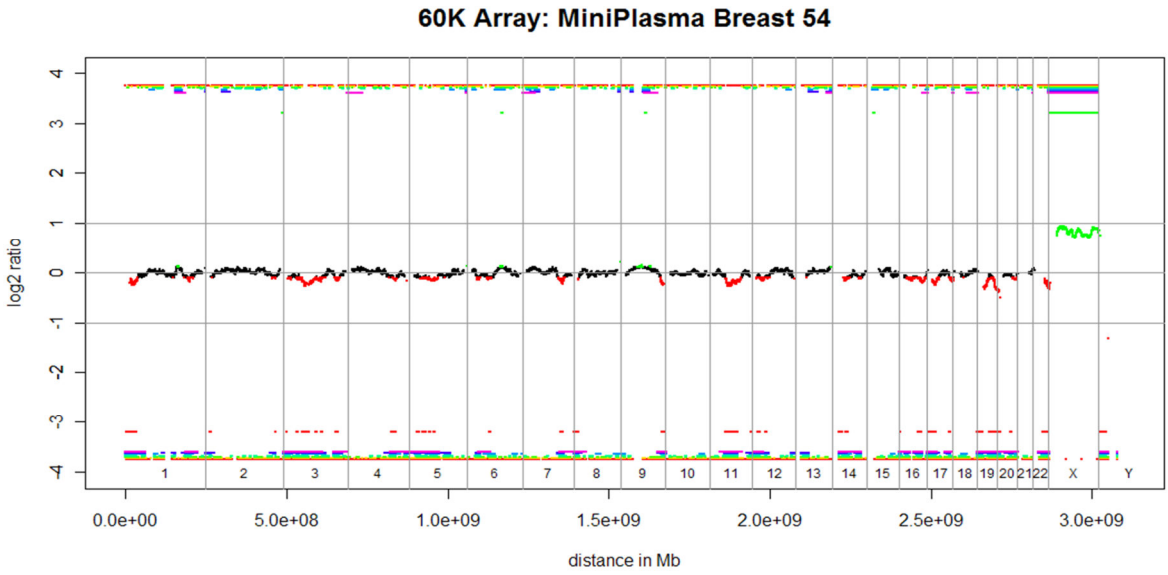

60K Array: MiniPlasma Breast 39
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

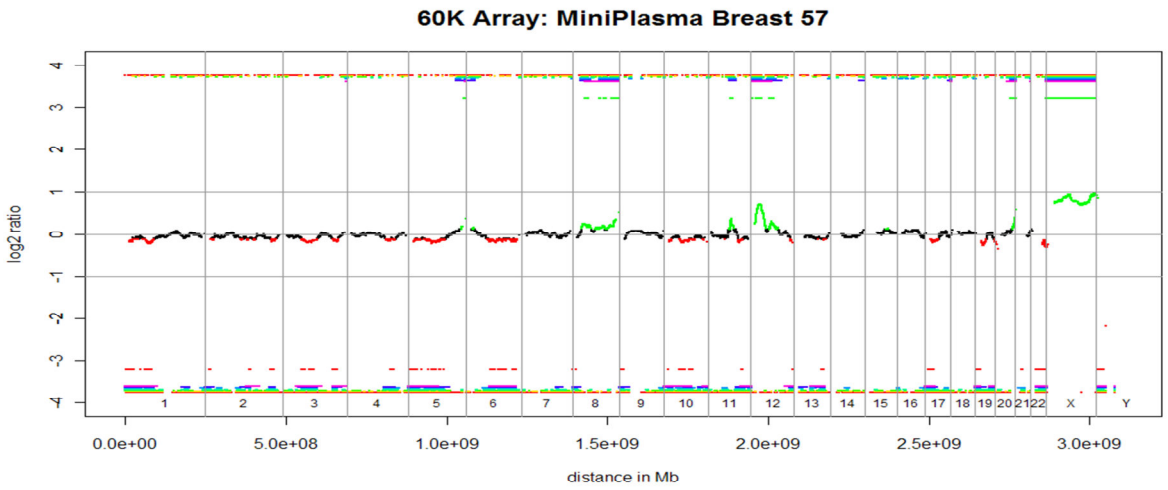

60K Array: MiniPlasma Breast 41
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

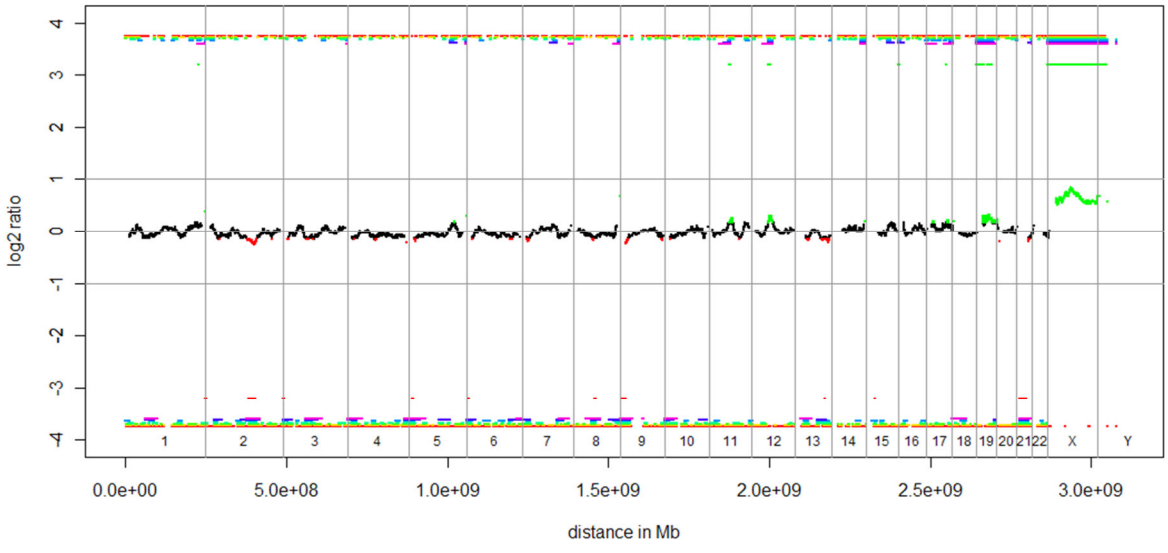

60K Array: MiniPlasma Breast 46
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

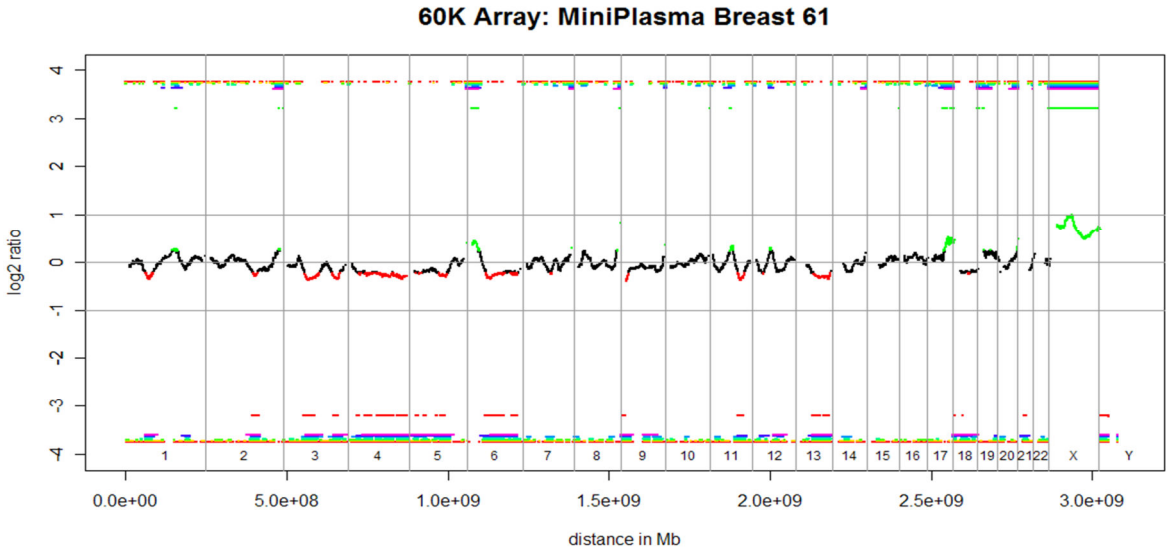

60K Array: MiniPlasma Breast 47
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

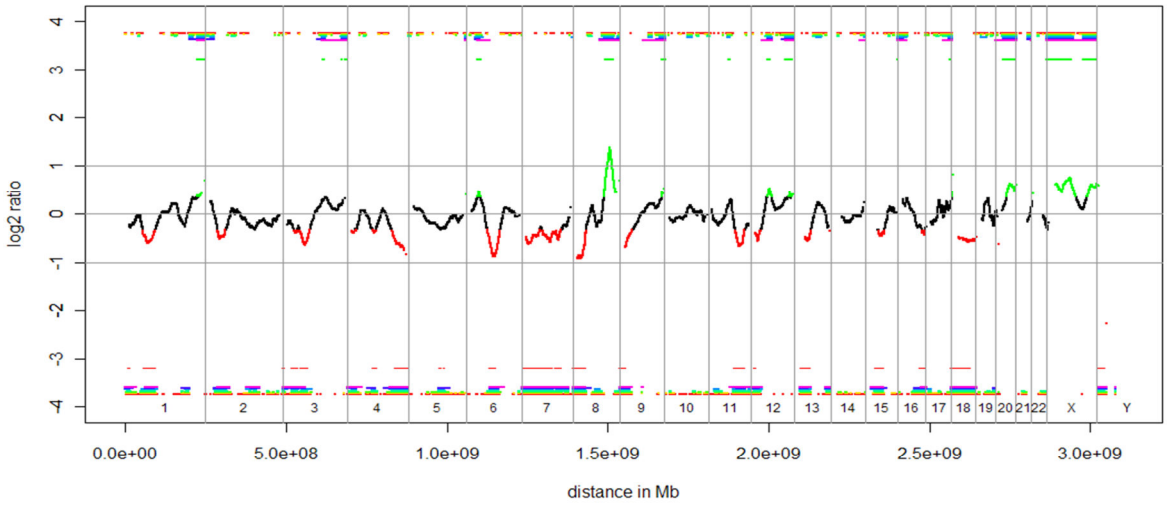

60K Array: MiniPlasma Breast 50
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

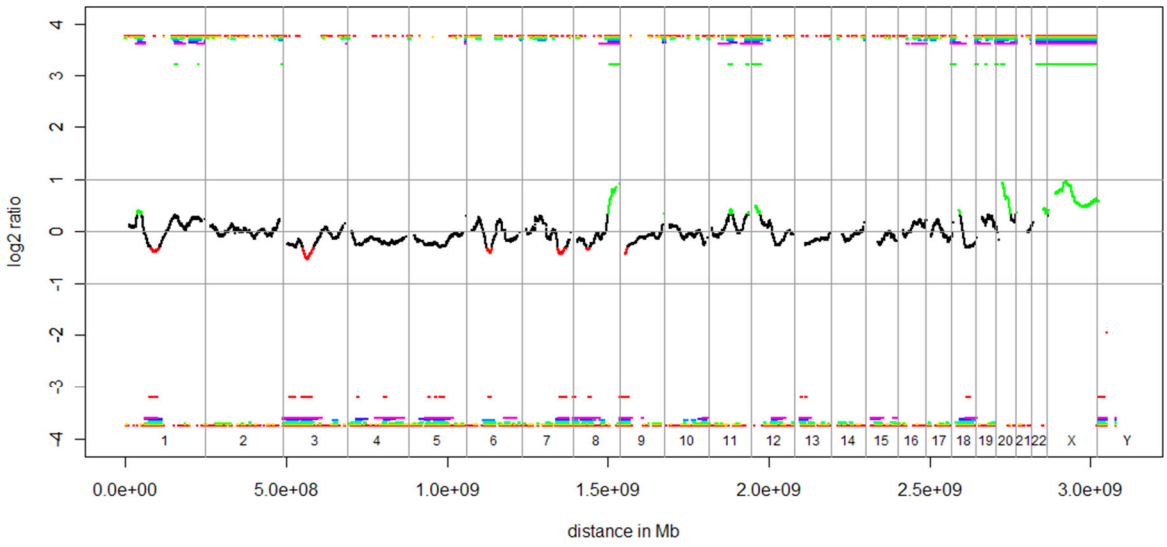

60K Array: MiniPlasma Breast 54
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

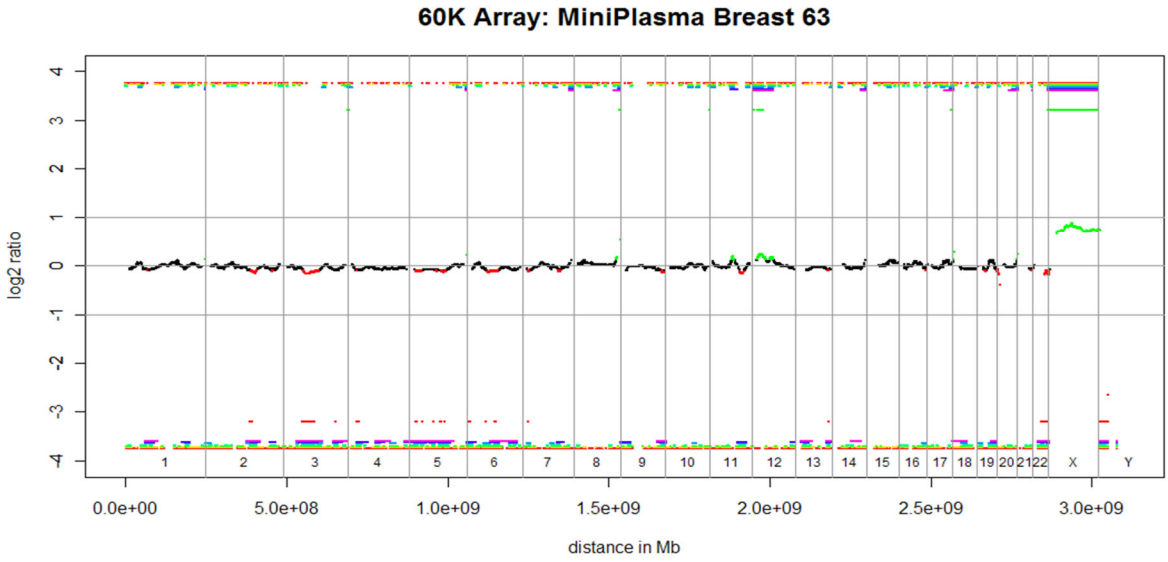

60K Array: MiniPlasma Breast 57
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

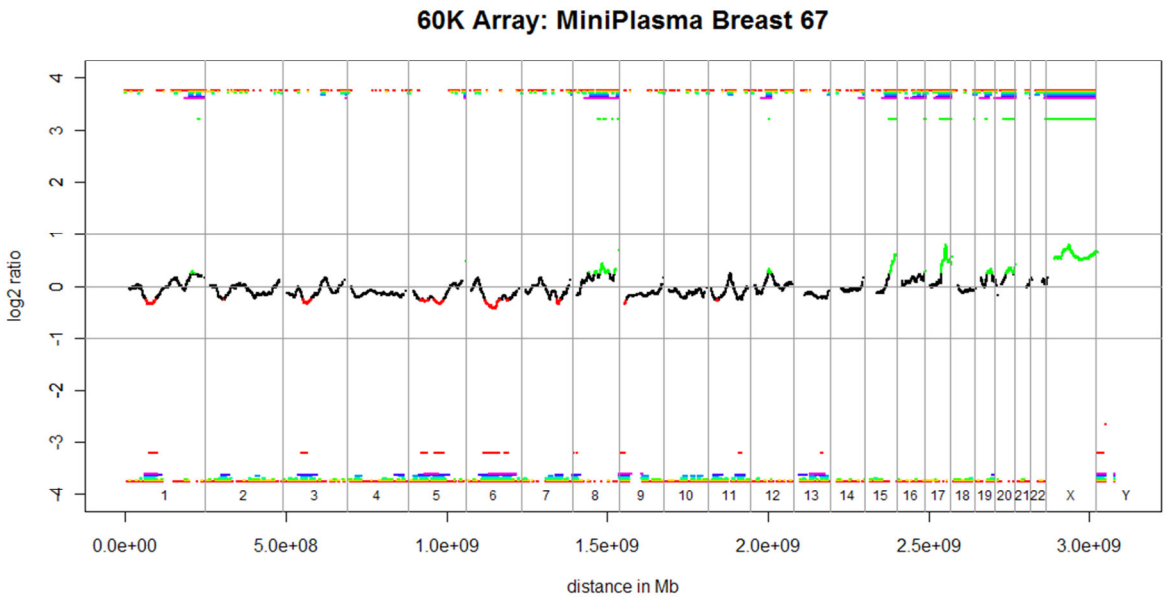

60K Array: MiniPlasma Breast 60
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

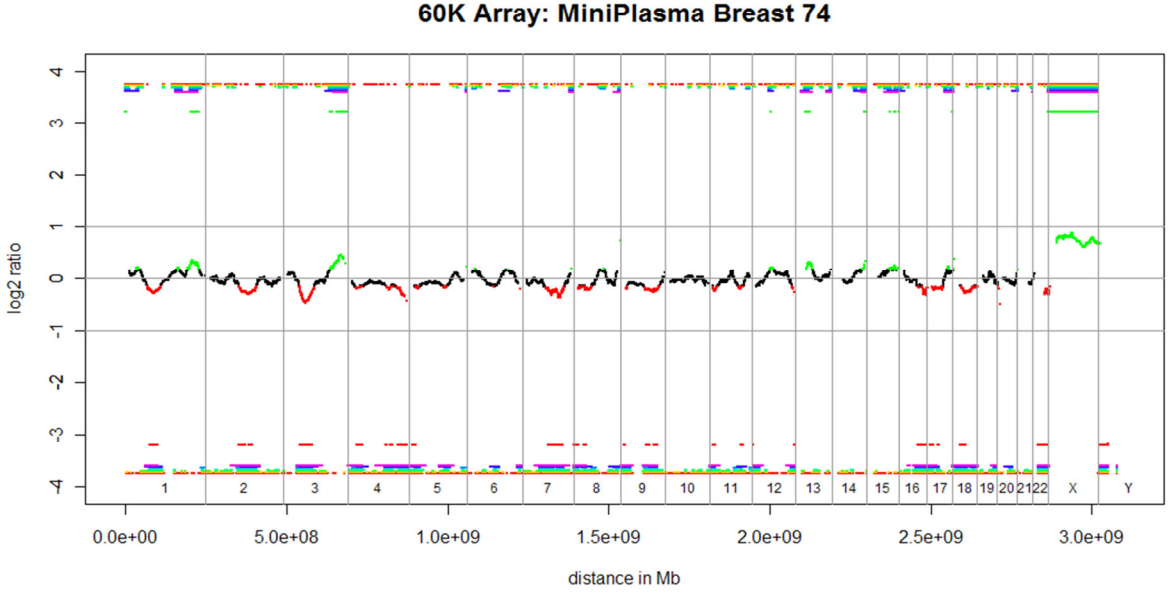

60K Array: MiniPlasma Breast 61
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

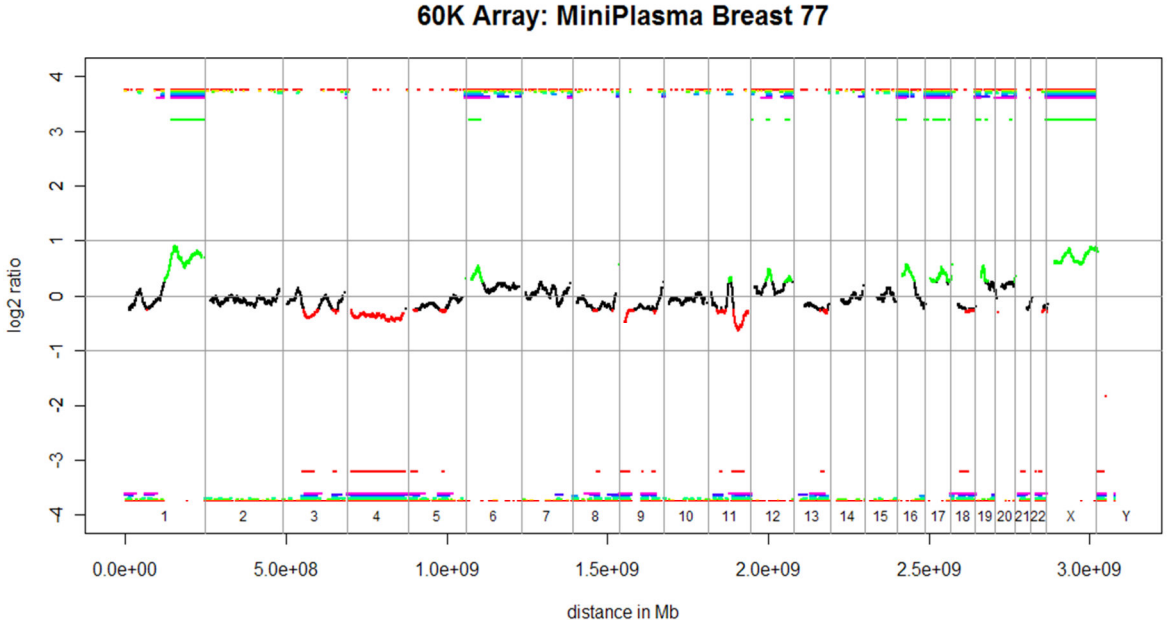

60K Array: MiniPlasma Breast 64
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

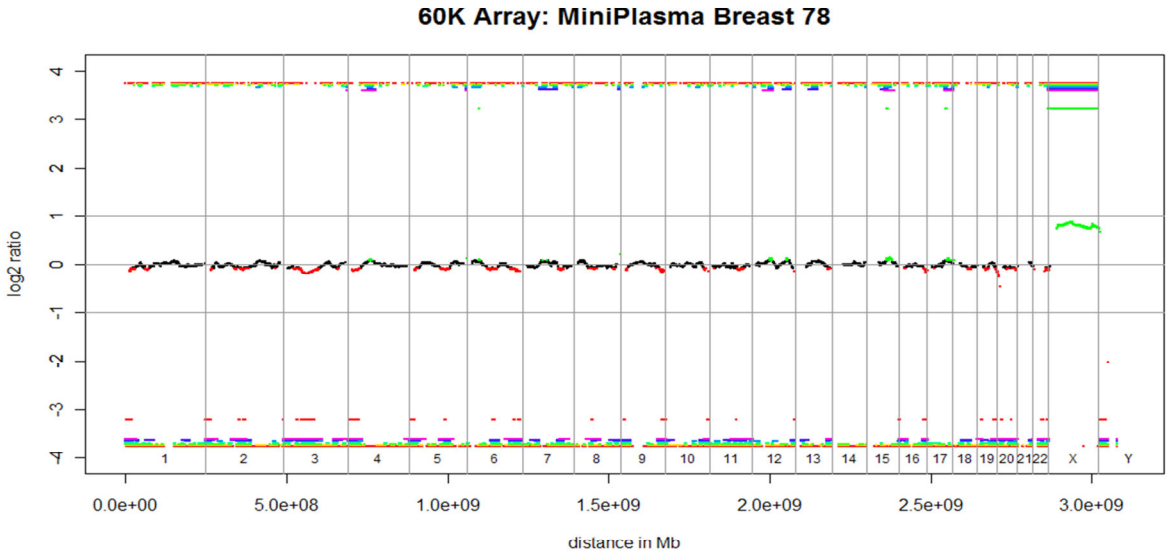

60K Array: MiniPlasma Breast 66
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

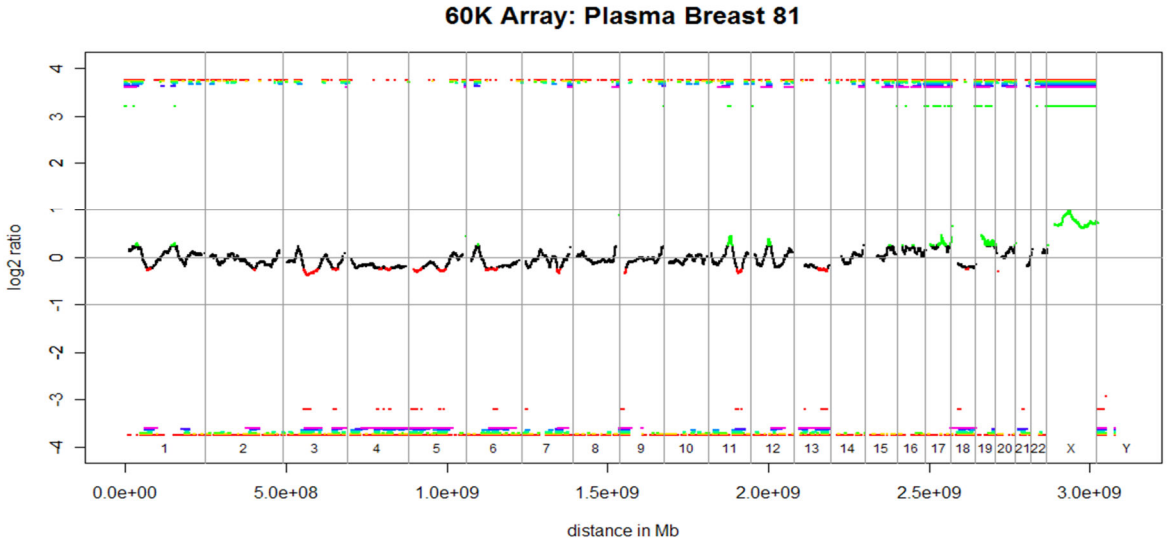

60K Array: MiniPlasma Breast 63
log2 ratio
distance in Mb


60K Array: MiniPlasma Breast 67
log2 ratio
distance in Mb


60K Array: MiniPlasma Breast 74
log2 ratio
distance in Mb


60K Array: MiniPlasma Breast 77
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb


60K Array: MiniPlasma Breast 78
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb


60K Array: Plasma Breast 81
log2 ratio
4
3
2
1
0
-1
-2
-3
-4
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
0.0e+00
5.0e+08
1.0e+09
1.5e+09
2.0e+09
2.5e+09
3.0e+09
distance in Mb

BREAST CANCER SAMPLES: Z-SCORES

Chromosome arm	B1			B2			B4		
	raw	ratio	Z-score	raw	ratio	Z-score	raw	ratio	Z-score
chr1p	210	0,0289	-3,88	1982	0,02983	-2,61	4660	0,02955	-2,99
chr1q	311	0,0428	8,32	2411	0,03628	0,44	7638	0,04844	15,14
chr10p	104	0,0143	1,48	921	0,01386	-0,40	2731	0,01732	13,95
chr10q	204	0,0281	-4,39	1923	0,02894	-1,09	3815	0,02419	-19,17
chr11p	130	0,0179	-2,04	1293	0,01946	1,17	2606	0,01653	-4,83
chr11q	148	0,0204	-2,72	1783	0,02683	2,20	4732	0,03001	4,62
chr12p	86	0,0118	-7,91	963	0,01449	1,75	2792	0,01771	13,43
chr12q	267	0,0367	2,66	2527	0,03803	4,29	5651	0,03583	1,52
chr13p	0	0,0000	0,00	0	0	0,00	0	0	0,00
chr13q	255	0,0351	-2,71	2358	0,03549	-2,16	4711	0,02987	-9,96
chr14p	0	0,0000	0,00	0	0	0,00	0	0	0,00
chr14q	188	0,0259	-1,45	1909	0,02873	2,01	3924	0,02488	-2,64
chr15p	0	0,0000	0,00	0	0	0,00	0	0	0,00
chr15q	171	0,0235	4,07	1335	0,02009	-2,70	2746	0,01741	-7,97
chr16p	82	0,0113	3,79	636	0,00957	-0,89	1311	0,00831	-4,32
chr16q	90	0,0124	3,04	780	0,01174	1,73	1844	0,01169	1,64
chr17p	41	0,0056	-2,60	329	0,00495	-4,67	675	0,00428	-6,68
chr17q	135	0,0186	8,41	764	0,0115	-1,39	1484	0,00941	-4,28
chr18p	29	0,0040	-6,74	333	0,00501	0,49	910	0,00577	5,87
chr18q	123	0,0169	-3,40	1355	0,02039	1,42	2882	0,01828	-1,52
chr19p	22	0,0030	1,46	200	0,00301	1,40	524	0,00332	2,47
chr19q	70	0,0096	5,27	495	0,00745	1,91	1352	0,00857	3,64
chr2p	200	0,0275	-6,91	1927	0,029	0,33	4632	0,02937	2,16
chr2q	407	0,0560	2,75	3656	0,05502	1,40	9185	0,05825	5,80
chr20p	84	0,0116	12,34	509	0,00766	-6,02	3043	0,0193	48,78
chr20q	78	0,0107	15,91	437	0,00658	3,45	995	0,00631	2,65
chr21p	9	0,0012	0,26	82	0,00123	0,24	240	0,00152	1,16
chr21q	102	0,0140	-0,52	931	0,01401	-0,58	1884	0,01195	-5,01
chr22p	0	0,0000	0,00	0	0	0,00	0	0	0,00
chr22q	33	0,0045	-2,41	322	0,00485	-1,42	974	0,00618	2,92
chr3p	219	0,0301	-0,53	1879	0,02828	-3,25	3614	0,02292	-11,07
chr3q	375	0,0516	8,83	3064	0,04611	1,94	7150	0,04534	0,98
chr4p	132	0,0182	0,13	1303	0,01961	2,35	2602	0,0165	-2,42
chr4q	487	0,0670	7,42	4019	0,06048	0,14	8564	0,05431	-6,74
chr5p	142	0,0195	1,09	1407	0,02117	5,72	3571	0,02264	9,90
chr5q	353	0,0486	-6,30	3396	0,05111	-0,23	7337	0,04653	-11,21
chr6p	127	0,0175	-2,37	1246	0,01875	0,09	3130	0,01985	2,20
chr6q	274	0,0377	-18,62	2972	0,04473	-2,49	7587	0,04811	5,30
chr7p	131	0,0180	-0,73	1202	0,01809	-0,62	5398	0,03423	29,21
chr7q	250	0,0344	-2,40	2340	0,03521	-0,11	4673	0,02963	-15,89
chr8p	91	0,0125	-8,96	945	0,01422	-1,61	1484	0,00941	-22,43
chr8q	430	0,0592	35,67	3178	0,04783	15,27	8697	0,05515	28,44
chr9p	96	0,0132	-8,23	1005	0,01512	-3,72	1994	0,01264	-9,56
chr9q	146	0,0201	-0,33	1291	0,01943	-2,40	2860	0,01814	-6,42
chrXp	122	0,0168	-15,74	1527	0,02298	-4,76	4124	0,02615	0,87
chrXq	304	0,0418	-11,70	3442	0,0518	-5,88	6861	0,04351	-10,72
chrYp	9	0,0012	2,86	44	0,00066	-1,05	103	0,00065	-1,11
chrYq	0	0,0000	-0,80	29	0,00044	70,27	5	3,2E-05	4,37

Chromosome arm	B9			B10			B15		
	raw	ratio	Z-score	*raw*	*ratio*	*Z-score*	raw	ratio	Z-score
chr1p	3356	0,02988	-2,55	2104	0,03024	-2,06	2278	0,03216	0,56
chr1q	3797	0,0338	-2,56	2521	0,03623	0,38	2683	0,03787	2,37
chr10p	1788	0,01592	8,13	972	0,01397	0,05	1008	0,01423	1,13
chr10q	3257	0,02899	-0,88	1995	0,02867	-2,12	2015	0,02844	-2,98
chr11p	2139	0,01904	0,32	1340	0,01926	0,76	1439	0,02031	2,92
chr11q	2431	0,02164	-1,75	1805	0,02594	1,52	1734	0,02448	0,41
chr12p	1648	0,01467	2,40	934	0,01342	-2,14	954	0,01347	-1,98
chr12q	4491	0,03998	6,76	2465	0,03542	1,00	2278	0,03216	-3,14
chr13p	0	0	0,00	0	0	0,00	0	0	0,00
chr13q	3842	0,0342	-3,94	2503	0,03597	-1,49	2406	0,03396	-4,28
chr14p	0	0	0,00	0	0	0,00	0	0	0,00
chr14q	3006	0,02676	-0,37	1964	0,02822	1,40	2172	0,03066	4,34
chr15p	0	0	0,00	0	0	0,00	0	0	0,00
chr15q	2342	0,02085	-1,21	1424	0,02046	-1,97	1358	0,01917	-4,52
chr16p	1034	0,0092	-1,88	637	0,00915	-2,02	553	0,00781	-5,70
chr16q	1283	0,01142	1,09	786	0,0113	0,83	802	0,01132	0,88
chr17p	659	0,00587	-1,93	376	0,0054	-3,32	383	0,00541	-3,31
chr17q	1367	0,01217	-0,46	812	0,01167	-1,16	809	0,01142	-1,50
chr18p	535	0,00476	-1,27	325	0,00467	-1,92	334	0,00471	-1,61
chr18q	2147	0,01911	-0,36	1401	0,02013	1,06	1432	0,02021	1,17
chr19p	377	0,00336	2,58	250	0,00359	3,39	252	0,00356	3,27
chr19q	1049	0,00934	4,82	491	0,00706	1,30	448	0,00632	0,18
chr2p	2926	0,02605	-14,14	1970	0,02831	-3,05	2139	0,03019	6,18
chr2q	5929	0,05278	-1,65	3935	0,05655	3,49	4422	0,06242	11,49
chr20p	820	0,0073	-7,72	681	0,00979	3,99	599	0,00846	-2,27
chr20q	1139	0,01014	14,13	410	0,00589	1,40	468	0,00661	3,54
chr21p	127	0,00113	-0,09	120	0,00172	1,81	100	0,00141	0,81
chr21q	1484	0,01321	-2,29	965	0,01387	-0,88	988	0,01395	-0,71
chr22p	0	0	0,00	0	0	0,00	0	0	0,00
chr22q	481	0,00428	-3,26	330	0,00474	-1,76	382	0,00539	0,36
chr3p	3187	0,02837	-3,11	2034	0,02923	-1,86	1962	0,0277	-4,10
chr3q	4647	0,04137	-4,01	2955	0,04246	-2,63	3191	0,04504	0,60
chr4p	2062	0,01836	0,42	1409	0,02025	3,33	1339	0,0189	1,26
chr4q	6976	0,0621	1,94	4332	0,06225	2,11	4116	0,0581	-2,52
chr5p	2207	0,01965	1,40	1400	0,02012	2,73	1454	0,02052	3,88
chr5q	5621	0,05004	-2,79	3591	0,0516	0,96	3716	0,05245	3,00
chr6p	2436	0,02169	5,73	1254	0,01802	-1,32	1280	0,01807	-1,23
chr6q	4351	0,03873	-16,26	3326	0,0478	4,57	3454	0,04876	6,78
chr7p	2117	0,01885	0,78	1225	0,0176	-1,52	1178	0,01663	-3,32
chr7q	4316	0,03842	8,96	2469	0,03548	0,65	2550	0,036	2,10
chr8p	1540	0,01371	-3,82	931	0,01338	-5,25	961	0,01357	-4,45
chr8q	6978	0,06212	40,97	2937	0,04221	5,17	2916	0,04116	3,29
chr9p	1399	0,01245	-10,01	1101	0,01582	-2,07	1271	0,01794	2,93
chr9q	2094	0,01864	-4,85	1383	0,01987	-1,01	1462	0,02064	1,37
chrXp	2667	0,02374	-3,41	1655	0,02378	-3,33	1377	0,01944	-11,04
chrXq	6190	0,0551	-3,94	3996	0,05742	-2,58	4082	0,05762	-2,47
chrYp	69	0,00061	-1,37	46	0,00066	-1,06	62	0,00088	0,39
chrYq	22	0,0002	31,09	27	0,00039	62,38	36	0,00051	81,95

Chromosome arm	B26			B30			B32		
	raw	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*	raw	ratio	Z-score
chr1p	31622	0,03043	-1,79	3676	0,03146	-0,38	1356	0,02756	-5,71
chr1q	37216	0,03582	-0,12	4350	0,03723	1,59	2375	0,04827	14,94
chr10p	15049	0,01448	2,19	1622	0,01388	-0,30	687	0,01396	0,03
chr10q	30270	0,02913	-0,36	3345	0,02863	-2,27	1409	0,02864	-2,24
chr11p	20147	0,01939	1,03	2216	0,01897	0,17	958	0,01947	1,19
chr11q	27193	0,02617	1,70	2987	0,02557	1,24	1247	0,02534	1,07
chr12p	13794	0,01327	-2,67	1617	0,01384	-0,62	605	0,0123	-6,23
chr12q	35525	0,03419	-0,57	4073	0,03486	0,28	1646	0,03345	-1,50
chr13p	0	0	0,00	0	0	0,00	0	0	0,00
chr13q	36554	0,03518	-2,59	4198	0,03593	-1,54	1736	0,03528	-2,44
chr14p	0	0	0,00	0	0	0,00	0	0	0,00
chr14q	29780	0,02866	1,92	3219	0,02755	0,58	1375	0,02794	1,06
chr15p	0	0	0,00	0	0	0,00	0	0	0,00
chr15q	20206	0,01945	-3,97	2475	0,02118	-0,55	973	0,01977	-3,32
chr16p	9281	0,00893	-2,63	1111	0,00951	-1,05	517	0,01051	1,67
chr16q	12237	0,01178	1,81	1376	0,01178	1,81	504	0,01024	-1,30
chr17p	5466	0,00526	-3,75	625	0,00535	-3,48	224	0,00455	-5,87
chr17q	11327	0,0109	-2,22	1322	0,01132	-1,65	537	0,01091	-2,20
chr18p	4797	0,00462	-2,30	564	0,00483	-0,81	193	0,00392	-7,22
chr18q	22039	0,02121	2,56	2533	0,02168	3,21	919	0,01868	-0,97
chr19p	3321	0,0032	2,04	392	0,00336	2,58	143	0,00291	1,05
chr19q	7203	0,00693	1,11	822	0,00704	1,27	379	0,0077	2,30
chr2p	29083	0,02799	-4,63	3340	0,02859	-1,69	1377	0,02799	-4,64
chr2q	58334	0,05614	2,93	6542	0,05599	2,73	2702	0,05491	1,26
chr20p	9227	0,00888	-0,27	1033	0,00884	-0,45	406	0,00825	-3,23
chr20q	6227	0,00599	1,70	656	0,00561	0,57	281	0,00571	0,85
chr21p	1216	0,00117	0,04	178	0,00152	1,17	49	0,001	-0,52
chr21q	15213	0,01464	0,78	1691	0,01447	0,42	722	0,01467	0,85
chr22p	0	0	0,00	0	0	0,00	0	0	0,00
chr22q	4789	0,00461	-2,19	577	0,00494	-1,12	262	0,00532	0,14
chr3p	30322	0,02918	-1,93	3487	0,02985	-0,96	1402	0,02849	-2,93
chr3q	44500	0,04282	-2,18	5091	0,04357	-1,24	2163	0,04396	-0,76
chr4p	19665	0,01892	1,30	2323	0,01988	2,77	990	0,02012	3,13
chr4q	64836	0,0624	2,27	7151	0,06121	0,94	3066	0,06231	2,18
chr5p	23287	0,02241	9,23	2368	0,02027	3,16	993	0,02018	2,91
chr5q	52700	0,05072	-1,17	5869	0,05023	-2,33	2559	0,05201	1,93
chr6p	20055	0,0193	1,14	2140	0,01832	-0,75	949	0,01929	1,12
chr6q	47301	0,04552	-0,66	5383	0,04607	0,61	2236	0,04544	-0,84
chr7p	19329	0,0186	0,33	2058	0,01761	-1,50	863	0,01754	-1,64
chr7q	38930	0,03746	6,26	4187	0,03584	1,65	1769	0,03595	1,98
chr8p	13437	0,01293	-7,19	1663	0,01423	-1,55	726	0,01475	0,70
chr8q	45640	0,04392	8,26	4824	0,04129	3,52	1983	0,0403	1,75
chr9p	16893	0,01626	-1,04	1925	0,01648	-0,53	807	0,0164	-0,71
chr9q	20127	0,01937	-2,58	2325	0,0199	-0,93	965	0,01961	-1,82
chrXp	24262	0,02335	-4,10	2831	0,02423	-2,54	1236	0,02512	-0,96
chrXq	59998	0,05774	-2,40	6577	0,05629	-3,25	2853	0,05798	-2,26
chrYp	655	0,00063	-1,27	68	0,00058	-1,59	45	0,00091	0,66
chrYq	60	5,8E-05	8,61	25	0,00021	34,05	17	0,00035	55,46

Chromosome arm	B36			B38			B39		
	raw	ratio	Z-score	*raw*	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*
chr1p	3543	0,03137	-0,51	8465	0,02811	-4,95	37865	0,03145	-0,40
chr1q	4166	0,03688	1,17	10587	0,03516	-0,91	50275	0,04176	7,06
chr10p	1615	0,0143	1,42	4699	0,01561	6,85	16043	0,01332	-2,62
chr10q	3397	0,03007	3,23	7264	0,02412	-19,42	34241	0,02844	-2,99
chr11p	2223	0,01968	1,63	4646	0,01543	-7,07	25581	0,02125	4,83
chr11q	2845	0,02519	0,95	9457	0,03141	5,69	28000	0,02326	-0,52
chr12p	1592	0,01409	0,30	2892	0,0096	-16,02	14444	0,012	-7,32
chr12q	3874	0,0343	-0,43	11307	0,03755	3,69	40910	0,03398	-0,83
chr13p	0	0	0,00	0	0	0,00	0	0	0,00
chr13q	3893	0,03446	-3,58	13973	0,04641	13,02	43020	0,03573	-1,82
chr14p	0	0	0,00	0	0	0,00	0	0	0,00
chr14q	3177	0,02813	1,28	10368	0,03443	8,91	33443	0,02778	0,86
chr15p	0	0	0,00	0	0	0,00	0	0	0,00
chr15q	2243	0,01986	-3,16	4519	0,01501	-12,71	24077	0,02	-2,88
chr16p	1064	0,00942	-1,30	3364	0,01117	3,48	11001	0,00914	-2,07
chr16q	1281	0,01134	0,92	4652	0,01545	9,24	13401	0,01113	0,50
chr17p	594	0,00526	-3,75	1366	0,00454	-5,92	5751	0,00478	-5,20
chr17q	1161	0,01028	-3,08	3661	0,01216	-0,48	14901	0,01238	-0,18
chr18p	498	0,00441	-3,77	1635	0,00543	3,46	5032	0,00418	-5,40
chr18q	2327	0,0206	1,71	7190	0,02388	6,27	23375	0,01941	0,06
chr19p	353	0,00313	1,80	772	0,00256	-0,12	3702	0,00307	1,62
chr19q	822	0,00728	1,64	2017	0,0067	0,75	9392	0,0078	2,45
chr2p	3169	0,02806	-4,30	7279	0,02417	-23,32	33975	0,02822	-3,50
chr2q	6272	0,05553	2,09	17705	0,0588	6,56	65182	0,05414	0,20
chr20p	944	0,00836	-2,74	3427	0,01138	11,51	9493	0,00788	-4,96
chr20q	698	0,00618	2,26	2277	0,00756	6,40	6757	0,00561	0,56
chr21p	152	0,00135	0,60	390	0,0013	0,44	1573	0,00131	0,47
chr21q	1589	0,01407	-0,45	3823	0,0127	-3,40	15565	0,01293	-2,90
chr22p	0	0	0,00	0	0	0,00	0	0	0,00
chr22q	562	0,00498	-1,00	1168	0,00388	-4,57	5708	0,00474	-1,76
chr3p	3316	0,02936	-1,67	6995	0,02323	-10,61	33592	0,0279	-3,80
chr3q	4933	0,04367	-1,12	16444	0,05461	12,60	57830	0,04803	4,35
chr4p	2142	0,01896	1,36	3108	0,01032	-11,91	22892	0,01901	1,43
chr4q	6962	0,06163	1,42	14447	0,04798	-13,79	74342	0,06175	1,55
chr5p	2380	0,02107	5,43	6675	0,02217	8,55	24624	0,02045	3,68
chr5q	6116	0,05414	7,05	11794	0,03917	-28,85	62757	0,05212	2,21
chr6p	2042	0,01808	-1,21	6608	0,02195	6,23	22158	0,0184	-0,58
chr6q	5256	0,04653	1,67	19398	0,06442	42,79	55171	0,04582	0,04
chr7p	1961	0,01736	-1,97	6073	0,02017	3,22	20746	0,01723	-2,20
chr7q	4056	0,03591	1,85	9012	0,02993	-15,05	43071	0,03577	1,47
chr8p	1555	0,01377	-3,58	3006	0,00998	-19,95	15634	0,01299	-6,96
chr8q	4549	0,04027	1,69	17866	0,05934	35,97	54232	0,04504	10,27
chr9p	1867	0,01653	-0,40	4206	0,01397	-6,44	20211	0,01679	0,20
chr9q	2215	0,01961	-1,83	5001	0,01661	-11,17	23903	0,01985	-1,07
chrXp	2800	0,02479	-1,55	7711	0,02561	-0,09	29600	0,02458	-1,91
chrXq	6685	0,05918	-1,56	13726	0,04559	-9,51	69802	0,05798	-2,26
chrYp	59	0,00052	-2,00	120	0,0004	-2,84	651	0,00054	-1,87
chrYq	8	7,1E-05	10,74	8	2,7E-05	3,53	66	5,5E-05	8,13

Chromosome arm	B41			B46			B47		
	raw	ratio	Z-score	raw	ratio	Z-score	raw	ratio	Z-score
chr1p	1083	0,02944	-3,14	1325	0,03058	-1,59	5444	0,03081	-1,27
chr1q	1289	0,03504	-1,06	1586	0,0366	0,83	6280	0,03554	-0,45
chr10p	463	0,01259	-5,68	640	0,01477	3,38	2202	0,01246	-6,20
chr10q	1035	0,02814	-4,15	1234	0,02848	-2,84	4703	0,02662	-9,94
chr11p	687	0,01868	-0,43	874	0,02017	2,63	3681	0,02083	3,98
chr11q	993	0,02699	2,33	767	0,0177	-4,75	5729	0,03242	6,46
chr12p	551	0,01498	3,52	603	0,01392	-0,34	2335	0,01321	-2,89
chr12q	1475	0,0401	6,91	1497	0,03455	-0,11	6419	0,03633	2,14
chr13p	0	0	0,00	0	0	0,00	0	0	0,00
chr13q	1171	0,03183	-7,24	1056	0,02437	-17,61	4594	0,026	-15,35
chr14p	0	0	0,00	0	0	0,00	0	0	0,00
chr14q	1104	0,03001	3,56	1214	0,02802	1,15	3775	0,02136	-6,90
chr15p	0	0	0,00	0	0	0,00	0	0	0,00
chr15q	823	0,02237	1,79	871	0,0201	-2,68	3212	0,01818	-6,47
chr16p	372	0,01011	0,59	467	0,01078	2,41	1733	0,00981	-0,24
chr16q	506	0,01376	5,81	320	0,00739	-7,08	2448	0,01385	6,01
chr17p	152	0,00413	-7,13	267	0,00616	-1,04	696	0,00394	-7,71
chr17q	630	0,01713	6,40	508	0,01172	-1,08	1878	0,01063	-2,60
chr18p	154	0,00419	-5,35	202	0,00466	-1,98	629	0,00356	-9,79
chr18q	650	0,01767	-2,37	973	0,02246	4,29	3026	0,01713	-3,13
chr19p	144	0,00391	4,49	130	0,003	1,37	603	0,00341	2,78
chr19q	384	0,01044	6,51	341	0,00787	2,56	1372	0,00776	2,39
chr2p	1011	0,02748	-7,10	1214	0,02802	-4,48	5490	0,03107	10,48
chr2q	1908	0,05187	-2,89	2483	0,05731	4,52	9686	0,05482	1,13
chr20p	333	0,00905	0,54	349	0,00805	-4,16	1599	0,00905	0,52
chr20q	537	0,0146	27,49	292	0,00674	3,94	1613	0,00913	11,10
chr21p	40	0,00109	-0,22	55	0,00127	0,36	157	0,00089	-0,86
chr21q	468	0,01272	-3,34	606	0,01399	-0,63	2435	0,01378	-1,07
chr22p	0	0	0,00	0	0	0,00	0	0	0,00
chr22q	132	0,00359	-5,52	250	0,00577	1,60	826	0,00467	-1,98
chr3p	1022	0,02778	-3,97	1293	0,02984	-0,96	5232	0,02961	-1,30
chr3q	1583	0,04303	-1,92	1934	0,04464	0,09	7979	0,04516	0,74
chr4p	653	0,01775	-0,50	772	0,01782	-0,40	3402	0,01925	1,80
chr4q	2427	0,06598	6,26	2638	0,06088	0,59	10973	0,0621	1,94
chr5p	909	0,02471	15,75	940	0,02169	7,20	3686	0,02086	4,84
chr5q	1938	0,05268	3,55	2381	0,05495	8,99	9122	0,05162	1,01
chr6p	696	0,01892	0,41	797	0,01839	-0,60	3388	0,01917	0,90
chr6q	1404	0,03817	-17,56	1955	0,04512	-1,58	8263	0,04676	2,20
chr7p	599	0,01628	-3,96	784	0,01809	-0,61	3214	0,01819	-0,43
chr7q	1149	0,03124	-11,36	1585	0,03658	3,76	6364	0,03602	2,16
chr8p	624	0,01696	10,26	483	0,01115	-14,91	2436	0,01379	-3,49
chr8q	1855	0,05043	19,95	2219	0,05121	21,37	9308	0,05268	24,00
chr9p	518	0,01408	-6,17	767	0,0177	2,36	2001	0,01132	-12,68
chr9q	623	0,01694	-10,15	925	0,02135	3,58	3641	0,02061	1,27
chrXp	830	0,02256	-5,50	1100	0,02539	-0,49	4436	0,0251	-0,99
chrXq	1832	0,0498	-7,04	2611	0,06026	-0,93	10614	0,06007	-1,04
chrYp	17	0,00046	-2,41	19	0,00044	-2,57	73	0,00041	-2,74
chrYq	11	0,0003	47,90	1	2,3E-05	2,96	1	5,7E-06	0,13

Chromosome arm	B49			B50			B54		
	raw	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*	raw	ratio	Z-score
chr1p	2191	0,02943	-3,15	28337	0,02815	-4,90	3007	0,02974	-2,74
chr1q	2958	0,03973	4,62	40059	0,03979	4,69	3647	0,03607	0,18
chr10p	1085	0,01457	2,57	14131	0,01404	0,34	1427	0,01411	0,65
chr10q	2245	0,03016	3,55	29645	0,02945	0,85	2953	0,0292	-0,08
chr11p	1303	0,0175	-2,83	19442	0,01931	0,87	1923	0,01902	0,27
chr11q	1398	0,01878	-3,93	25862	0,02569	1,33	2635	0,02606	1,61
chr12p	814	0,01093	-11,18	12116	0,01204	-7,18	1386	0,01371	-1,10
chr12q	2350	0,03157	-3,89	34739	0,03451	-0,16	3639	0,03599	1,71
chr13p	0	0	0	0	0	0,00	0	0	0,00
chr13q	1966	0,02641	-14,78	37056	0,03681	-0,32	3617	0,03577	-1,76
chr14p	0	0	0	0	0	0,00	0	0	0,00
chr14q	2371	0,03185	5,78	28237	0,02805	1,19	2743	0,02713	0,07
chr15p	0	0	0	0	0	0,00	0	0	0,00
chr15q	1544	0,02074	-1,42	21089	0,02095	-1,01	2182	0,02158	0,23
chr16p	970	0,01303	8,55	9647	0,00958	-0,85	978	0,00967	-0,61
chr16q	695	0,00934	-3,13	10607	0,01054	-0,70	1124	0,01112	0,47
chr17p	264	0,00355	-8,88	5499	0,00546	-3,14	564	0,00558	-2,80
chr17q	682	0,00916	-4,63	11472	0,0114	-1,53	1055	0,01043	-2,87
chr18p	345	0,00463	-2,18	4458	0,00443	-3,63	459	0,00454	-2,85
chr18q	1698	0,02281	4,78	18728	0,0186	-1,07	2045	0,02022	1,19
chr19p	184	0,00247	-0,43	3361	0,00334	2,52	345	0,00341	2,77
chr19q	424	0,0057	-0,79	7144	0,0071	1,37	689	0,00681	0,93
chr2p	1892	0,02542	-17,24	28939	0,02875	-0,91	2963	0,0293	1,81
chr2q	3671	0,04931	-6,37	56587	0,05621	3,03	5261	0,05203	-2,67
chr20p	672	0,00903	0,42	8523	0,00847	-2,22	794	0,00785	-5,11
chr20q	497	0,00668	3,75	5639	0,0056	0,53	573	0,00567	0,72
chr21p	54	0,00073	-1,38	1311	0,0013	0,46	141	0,00139	0,75
chr21q	1058	0,01421	-0,14	14777	0,01468	0,86	1440	0,01424	-0,08
chr22p	0	0	0	0	0	0,00	0	0	0,00
chr22q	391	0,00525	-0,09	4609	0,00458	-2,29	459	0,00454	-2,42
chr3p	1953	0,02623	-6,23	30765	0,03056	0,09	2882	0,0285	-2,92
chr3q	4491	0,06033	19,77	45407	0,04511	0,68	4342	0,04294	-2,04
chr4p	1385	0,0186	0,81	18085	0,01797	-0,18	1862	0,01841	0,51
chr4q	3898	0,05236	-8,91	56200	0,05583	-5,05	6339	0,06269	2,60
chr5p	1561	0,02097	5,14	21016	0,02088	4,88	1951	0,01929	0,40
chr5q	3733	0,05015	-2,54	54850	0,05449	7,87	5346	0,05287	3,99
chr6p	1285	0,01726	-2,78	19977	0,01984	2,19	1871	0,0185	-0,39
chr6q	3689	0,04955	8,61	47365	0,04705	2,86	4515	0,04465	-2,66
chr7p	1561	0,02097	4,70	18447	0,01832	-0,18	1790	0,0177	-1,33
chr7q	2601	0,03494	-0,88	36793	0,03655	3,67	3691	0,0365	3,53
chr8p	677	0,00909	-23,80	14459	0,01436	-0,99	1352	0,01337	-5,29
chr8q	4714	0,06332	43,14	41989	0,04171	4,28	4337	0,04289	6,40
chr9p	1356	0,01822	3,57	17304	0,01719	1,15	1642	0,01624	-1,09
chr9q	1342	0,01803	-6,76	19886	0,01975	-1,38	2102	0,02079	1,83
chrXp	1525	0,02049	-9,18	23576	0,02342	-3,98	2717	0,02687	2,14
chrXq	4909	0,06594	2,40	57376	0,057	-2,83	6268	0,06199	0,08
chrYp	40	0,00054	-1,90	812	0,00081	-0,07	60	0,00059	-1,52
chrYq	2	2,7E-05	3,58	337	0,00033	53,71	3	3E-05	4,03

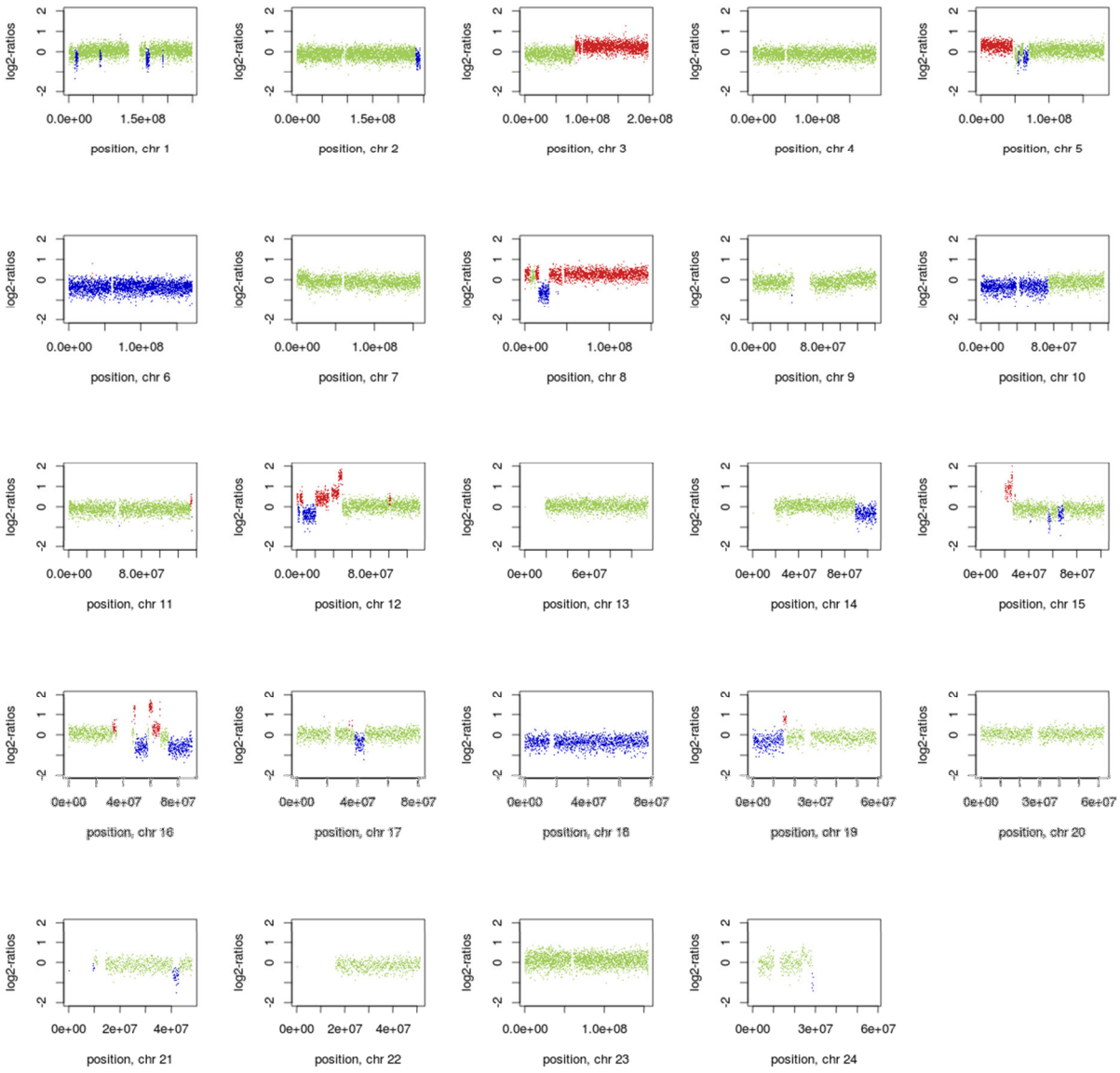
Chromosome arm	B60			B61			B63		
	raw	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*
chr1p	19006	0,03052	-1,67	3378	0,03057	-1,59	15756	0,03131	-0,59
chr1q	23119	0,03712	1,46	4158	0,03763	2,08	18346	0,03646	0,65
chr10p	8495	0,01364	-1,30	1585	0,01435	1,62	6911	0,01373	-0,92
chr10q	18149	0,02914	-0,31	3050	0,02761	-6,17	13867	0,02756	-6,36
chr11p	11723	0,01882	-0,13	2126	0,01924	0,73	9672	0,01922	0,68
chr11q	15653	0,02514	0,91	2758	0,02496	0,78	12913	0,02566	1,31
chr12p	8636	0,01387	-0,52	1566	0,01417	0,60	9368	0,01862	16,74
chr12q	21365	0,03431	-0,42	3778	0,03419	-0,56	17376	0,03453	-0,14
chr13p	0	0	0,00	0	0	0,00	0,00	0,00	0,00
chr13q	22600	0,03629	-1,04	3946	0,03572	-1,84	17859	0,03549	-2,16
chr14p	0	0	0,00	0	0	0,00	0,00	0,00	0,00
chr14q	17199	0,02762	0,66	3062	0,02771	0,78	13760	0,02734	0,33
chr15p	0	0	0,00	0	0	0,00	0,00	0,00	0,00
chr15q	13177	0,02116	-0,60	2271	0,02056	-1,79	10333	0,02053	-1,83
chr16p	5631	0,00904	-2,33	1082	0,00979	-0,28	4689	0,00932	-1,58
chr16q	7020	0,01127	0,79	1203	0,01089	0,01	5799	0,01152	1,29
chr17p	3475	0,00558	-2,79	585	0,00529	-3,64	2662	0,00529	-3,66
chr17q	7147	0,01148	-1,42	1352	0,01224	-0,37	5684	0,01129	-1,67
chr18p	3123	0,00501	0,52	511	0,00463	-2,24	2355	0,00468	-1,86
chr18q	12890	0,0207	1,85	2224	0,02013	1,06	10212	0,02029	1,28
chr19p	2010	0,00323	2,15	384	0,00348	2,99	1539	0,00306	1,57
chr19q	4373	0,00702	1,25	789	0,00714	1,44	3731	0,00741	1,85
chr2p	17374	0,0279	-5,06	3322	0,03007	5,56	14035	0,02789	-5,11
chr2q	34071	0,05471	0,98	6056	0,05481	1,12	27012	0,05368	-0,43
chr20p	5203	0,00836	-2,75	866	0,00784	-5,18	4298	0,00854	-1,87
chr20q	3526	0,00566	0,71	703	0,00636	2,81	2968	0,0059	1,41
chr21p	750	0,0012	0,15	153	0,00138	0,72	675	0,00134	0,59
chr21q	8700	0,01397	-0,66	1673	0,01514	1,86	7108	0,01412	-0,33
chr22p	0	0	0,00	0	0	0,00	0,00	0,00	0,00
chr22q	2857	0,00459	-2,26	466	0,00422	-3,47	2159	0,00429	-3,23
chr3p	18667	0,02998	-0,77	3345	0,03028	-0,33	14618	0,02905	-2,12
chr3q	27635	0,04438	-0,23	4723	0,04275	-2,27	21856	0,04343	-1,42
chr4p	12158	0,01952	2,22	2110	0,0191	1,56	9746	0,01937	1,97
chr4q	39304	0,06311	3,07	6479	0,05864	-1,91	31058	0,06172	1,51
chr5p	12728	0,02044	3,64	2285	0,02068	4,33	10210	0,02029	3,21
chr5q	32212	0,05173	1,25	5764	0,05217	2,32	26301	0,05226	2,54
chr6p	11741	0,01885	0,28	2386	0,0216	5,56	9210	0,0183	-0,78
chr6q	28544	0,04584	0,07	4952	0,04482	-2,27	22715	0,04514	-1,54
chr7p	10912	0,01752	-1,67	1988	0,01799	-0,80	8879	0,01764	-1,44
chr7q	22501	0,03613	2,49	4034	0,03651	3,56	17660	0,03509	-0,45
chr8p	8519	0,01368	-3,95	1540	0,01394	-2,83	7149	0,01421	-1,67
chr8q	26700	0,04288	6,37	4854	0,04393	8,28	22389	0,04449	9,28
chr9p	10276	0,0165	-0,47	1685	0,01525	-3,42	8137	0,01617	-1,25
chr9q	12106	0,01944	-2,36	2062	0,01866	-4,78	9495	0,01887	-4,14
chrXp	15119	0,02428	-2,46	2944	0,02665	1,74	12311	0,02446	-2,13
chrXq	35771	0,05744	-2,57	6221	0,05631	-3,24	30061	0,05974	-1,23
chrYp	439	0,0007	-0,76	65	0,00059	-1,55	327	0,00065	-1,13
chrYq	133	0,00021	33,98	0	0	-0,80	30	6E-05	8,91

Chromosome arm	B64			B66			B67		
	raw	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*
chr1p	3327	0,02471	-9,60	3387	0,02845	-4,49	1508	0,028	-5,11
chr1q	5506	0,0409	6,02	5176	0,04348	9,15	2346	0,04356	9,25
chr10p	1851	0,01375	-0,86	2063	0,01733	14,00	740	0,01374	-0,89
chr10q	3954	0,02937	0,55	2942	0,02471	-17,18	1409	0,02616	-11,66
chr11p	2502	0,01858	-0,62	2138	0,01796	-1,89	864	0,01604	-5,82
chr11q	2654	0,01971	-3,22	3599	0,03023	4,79	1241	0,02304	-0,68
chr12p	1476	0,01096	-11,08	2483	0,02086	24,90	680	0,01263	-5,03
chr12q	5594	0,04155	8,75	3654	0,0307	-4,99	2076	0,03855	4,95
chr13p	0,00	0,00	0,00	0,00	0	0,00	0,00	0,00	0,00
chr13q	5162	0,03834	1,81	3884	0,03263	-6,13	1782	0,03309	-5,49
chr14p	0,00	0,00	0,00	0,00	0	0,00	0,00	0,00	0,00
chr14q	3495	0,02596	-1,34	3007	0,02526	-2,19	1474	0,02737	0,36
chr15p	0,00	0,00	0,00	0,00	0	0,00	0,00	0,00	0,00
chr15q	2218	0,01647	-9,82	1873	0,01573	-11,28	1237	0,02297	2,97
chr16p	1561	0,01159	4,63	957	0,00804	-5,06	485	0,00901	-2,43
chr16q	1083	0,00804	-5,75	1314	0,01104	0,31	631	0,01172	1,69
chr17p	582	0,00432	-6,56	594	0,00499	-4,56	238	0,00442	-6,27
chr17q	1616	0,012	-0,69	1096	0,00921	-4,56	970	0,01801	7,62
chr18p	506	0,00376	-8,38	1214	0,0102	37,23	295	0,00548	3,80
chr18q	1952	0,0145	-6,78	1797	0,0151	-5,95	1068	0,01983	0,64
chr19p	430	0,00319	2,03	375	0,00315	1,88	148	0,00275	0,51
chr19q	745	0,00553	-1,04	987	0,00829	3,20	422	0,00784	2,50
chr2p	3706	0,02753	-6,89	3585	0,03012	5,80	1493	0,02772	-5,93
chr2q	6920	0,0514	-3,53	6826	0,05734	4,57	3074	0,05708	4,21
chr20p	1230	0,00914	0,93	2244	0,01885	46,68	491	0,00912	0,84
chr20q	1125	0,00836	8,78	763	0,00641	2,95	367	0,00681	4,16
chr21p	116	0,00086	-0,94	144	0,00121	0,17	72	0,00134	0,57
chr21q	2014	0,01496	1,46	1597	0,01342	-1,85	700	0,013	-2,75
chr22p	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
chr22q	464	0,00345	-5,98	729	0,00612	2,75	248	0,00461	-2,20
chr3p	3233	0,02401	-9,47	2707	0,02274	-11,33	1563	0,02902	-2,16
chr3q	7661	0,0569	15,47	5623	0,04724	3,35	2709	0,0503	7,20
chr4p	2479	0,01841	0,51	1932	0,01623	-2,84	958	0,01779	-0,45
chr4q	7855	0,05834	-2,25	6518	0,05475	-6,24	3416	0,06343	3,42
chr5p	3441	0,02556	18,16	2565	0,02155	6,78	1148	0,02132	6,13
chr5q	6862	0,05097	-0,57	5531	0,04646	-11,36	2702	0,05017	-2,47
chr6p	2896	0,02151	5,39	2228	0,01872	0,02	970	0,01801	-1,34
chr6q	5954	0,04422	-3,64	5726	0,0481	5,27	2068	0,0384	-17,03
chr7p	1642	0,0122	-11,51	3938	0,03308	27,09	970	0,01801	-0,76
chr7q	3893	0,02892	-17,92	3603	0,03027	-14,09	1785	0,03315	-5,96
chr8p	735	0,00546	-39,53	1141	0,00958	-21,67	598	0,0111	-15,10
chr8q	11102	0,08246	77,54	6903	0,05799	33,54	3327	0,06178	40,36
chr9p	1566	0,01163	-11,95	1499	0,01259	-9,69	749	0,01391	-6,58
chr9q	3216	0,02389	11,48	2077	0,01745	-8,56	932	0,01731	-9,00
chrXp	3485	0,02588	0,39	3227	0,02711	2,56	1168	0,02169	-7,05
chrXq	6735	0,05002	-6,91	5298	0,04451	-10,14	2701	0,05015	-6,84
chrYp	89	0,00066	-1,06	97	0,00081	-0,02	27	0,0005	-2,14
chrYq	2	1,5E-05	1,62	0	0	-0,80	4	7,4E-05	11,30

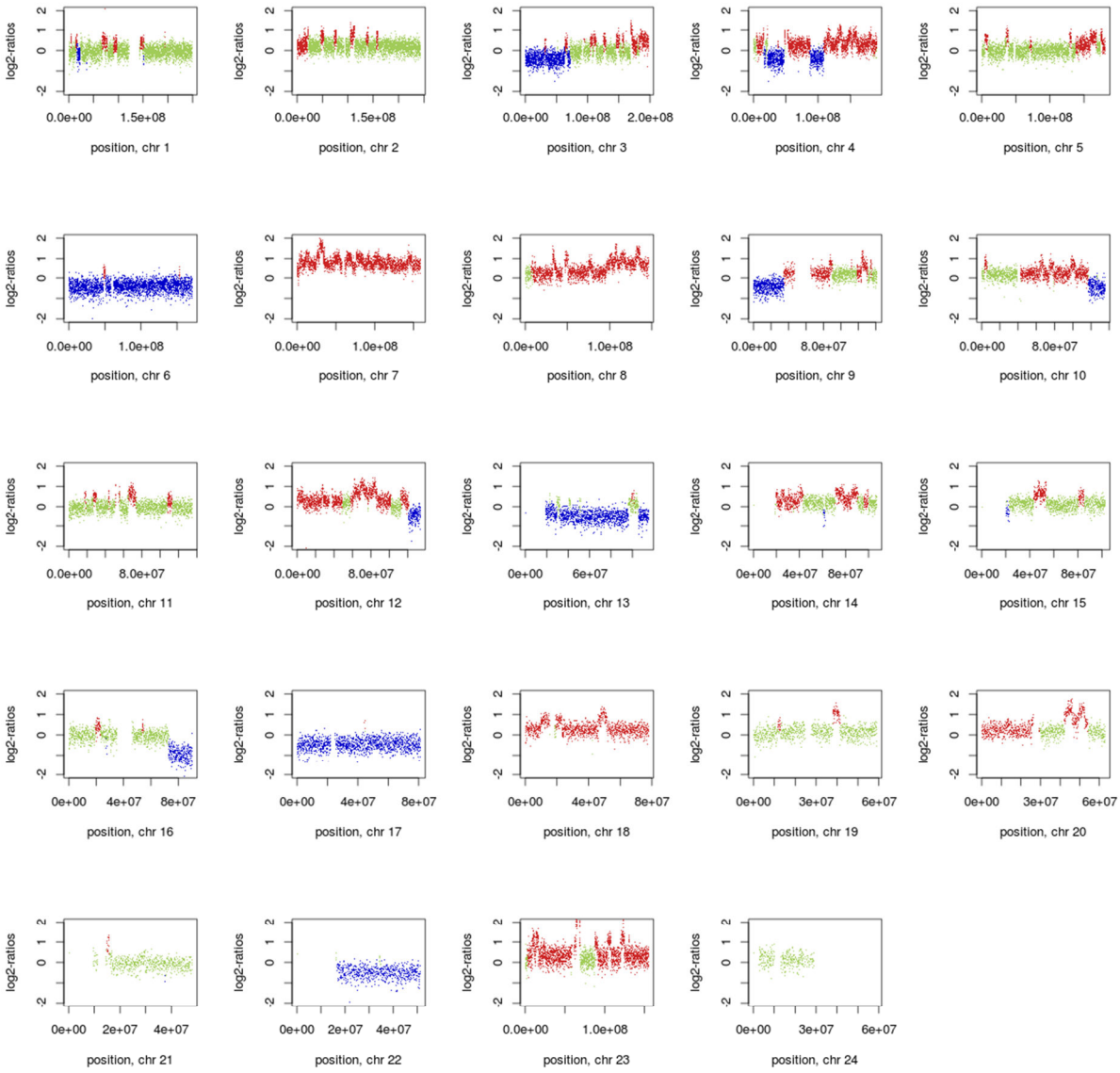
Chromosome arm	B74			B78			B81		
	raw	*ratio*	*Z-score*	raw	ratio	Z-score	raw	ratio	Z-score
chr1p	30183	0,02905	-3,68	4984	0,03143	-0,43	40811	0,03106	-0,93
chr1q	41447	0,03989	4,80	5863	0,03697	1,28	48739	0,0371	1,43
chr10p	13955	0,01343	-2,18	2265	0,01428	1,36	17804	0,01355	-1,68
chr10q	29802	0,02868	-2,07	4621	0,02914	-0,33	37684	0,02868	-2,06
chr11p	19082	0,01836	-1,07	2904	0,01831	-1,17	25264	0,01923	0,70
chr11q	26163	0,02518	0,94	4023	0,02537	1,09	32935	0,02507	0,86
chr12p	12964	0,01248	-5,58	2224	0,01402	0,05	17722	0,01349	-1,90
chr12q	36880	0,03549	1,08	5318	0,03353	-1,40	44965	0,03422	-0,52
chr13p	0	0	0,00	0,00	0,00	0,00	0,00	0,00	0,00
chr13q	42672	0,04107	5,60	5739	0,03619	-1,18	47021	0,03579	-1,74
chr14p	0	0	0,00	0,00	0,00	0,00	0,00	0,00	0,00
chr14q	29839	0,02872	1,99	4405	0,02778	0,86	36430	0,02773	0,80
chr15p	0	0	0,00	0,00	0,00	0,00	0,00	0,00	0,00
chr15q	22466	0,02162	0,31	3288	0,02073	-1,44	27252	0,02074	-1,42
chr16p	10102	0,00972	-0,47	1451	0,00915	-2,03	11958	0,0091	-2,17
chr16q	9986	0,00961	-2,58	1765	0,01113	0,50	15080	0,01148	1,20
chr17p	4791	0,00461	-5,69	818	0,00516	-4,05	7099	0,0054	-3,32
chr17q	11321	0,0109	-2,23	1747	0,01102	-2,06	15028	0,01144	-1,48
chr18p	4924	0,00474	-1,44	776	0,00489	-0,34	6183	0,00471	-1,67
chr18q	19982	0,01923	-0,20	3247	0,02047	1,54	26954	0,02052	1,59
chr19p	3201	0,00308	1,64	487	0,00307	1,61	4334	0,0033	2,39
chr19q	7194	0,00692	1,10	1144	0,00721	1,55	9339	0,00711	1,38
chr2p	29042	0,02795	-4,82	4511	0,02845	-2,39	37127	0,02826	-3,30
chr2q	53579	0,05156	-3,31	8669	0,05467	0,92	71694	0,05457	0,79
chr20p	8725	0,0084	-2,55	1303	0,00822	-3,40	11125	0,00847	-2,21
chr20q	5630	0,00542	-0,02	865	0,00545	0,09	7460	0,00568	0,76
chr21p	1377	0,00133	0,53	206	0,0013	0,45	1493	0,00114	-0,07
chr21q	14090	0,01356	-1,54	2231	0,01407	-0,45	18733	0,01426	-0,04
chr22p	0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
chr22q	4048	0,0039	-4,52	764	0,00482	-1,51	6371	0,00485	-1,41
chr3p	30432	0,02929	-1,77	4599	0,029	-2,19	39558	0,03011	-0,57
chr3q	56003	0,0539	11,70	6879	0,04338	-1,49	58620	0,04462	0,07
chr4p	18788	0,01808	0,00	2981	0,0188	1,10	26030	0,01981	2,66
chr4q	61226	0,05892	-1,60	10047	0,06335	3,34	81855	0,0623	2,17
chr5p	21177	0,02038	3,47	3288	0,02073	4,48	26680	0,02031	3,27
chr5q	53221	0,05122	0,04	8291	0,05228	2,58	68442	0,05209	2,14
chr6p	21150	0,02035	3,17	3036	0,01914	0,84	25426	0,01935	1,24
chr6q	51643	0,0497	8,95	7282	0,04592	0,26	59841	0,04555	-0,60
chr7p	18446	0,01775	-1,24	2849	0,01797	-0,85	23156	0,01763	-1,48
chr7q	33587	0,03232	-8,28	5690	0,03588	1,78	46732	0,03557	0,90
chr8p	12229	0,01177	-12,22	2263	0,01427	-1,40	18194	0,01385	-3,22
chr8q	44859	0,04317	6,91	6662	0,04201	4,82	56769	0,04321	6,97
chr9p	17011	0,01637	-0,78	2700	0,01703	0,77	21663	0,01649	-0,50
chr9q	18350	0,01766	-7,90	3037	0,01915	-3,26	25675	0,01954	-2,04
chrXp	26563	0,02556	-0,18	3849	0,02427	-2,47	31459	0,02394	-3,05
chrXq	59958	0,0577	-2,42	9343	0,05892	-1,71	75924	0,05779	-2,37
chrYp	753	0,00072	-0,63	133	0,00084	0,15	911	0,00069	-0,84
chrYq	258	0,00025	39,63	37	0,00023	37,19	273	0,00021	33,04

Chromosome arm	B83		
	raw	ratio	Z-score
chr1p	9838	0,02551	-8,51
chr1q	18022	0,04673	13,08
chr10p	9351	0,02425	42,69
chr10q	8413	0,02181	-28,22
chr11p	6790	0,01761	-2,62
chr11q	7183	0,01863	-4,05
chr12p	8871	0,023	32,69
chr12q	10859	0,02816	-8,20
chr13p	0	0,00	0,00
chr13q	13027	0,03378	-4,53
chr14p	0	0,00	0,00
chr14q	7617	0,01975	-8,85
chr15p	0	0,00	0,00
chr15q	7165	0,01858	-5,68
chr16p	2929	0,00759	-6,28
chr16q	3780	0,0098	-2,19
chr17p	3430	0,00889	7,14
chr17q	5070	0,01315	0,89
chr18p	2385	0,00618	8,80
chr18q	8448	0,02191	3,53
chr19p	1480	0,00384	4,22
chr19q	2847	0,00738	1,81
chr2p	10571	0,02741	-7,46
chr2q	15543	0,0403	-18,65
chr20p	2243	0,00582	-14,70
chr20q	2132	0,00553	0,31
chr21p	418	0,00108	-0,24
chr21q	3979	0,01032	-8,51
chr22p	0	0,00	0,00
chr22q	2106	0,00546	0,59
chr3p	16212	0,04204	16,83
chr3q	17591	0,04561	1,32
chr4p	5533	0,01435	-5,73
chr4q	17669	0,04582	-16,20
chr5p	12230	0,03171	35,61
chr5q	14788	0,03835	-30,82
chr6p	9530	0,02471	11,55
chr6q	22641	0,05871	29,66
chr7p	5534	0,01435	-7,53
chr7q	13282	0,03444	-2,29
chr8p	3513	0,00911	-23,73
chr8q	21787	0,05649	30,86
chr9p	10238	0,02655	23,22
chr9q	5734	0,01487	-16,59
chrXp	10193	0,02643	1,36
chrXq	24407	0,06329	0,84
chrYp	273	0,00071	-0,74
chrYq	1	2,6E-06	-0,37

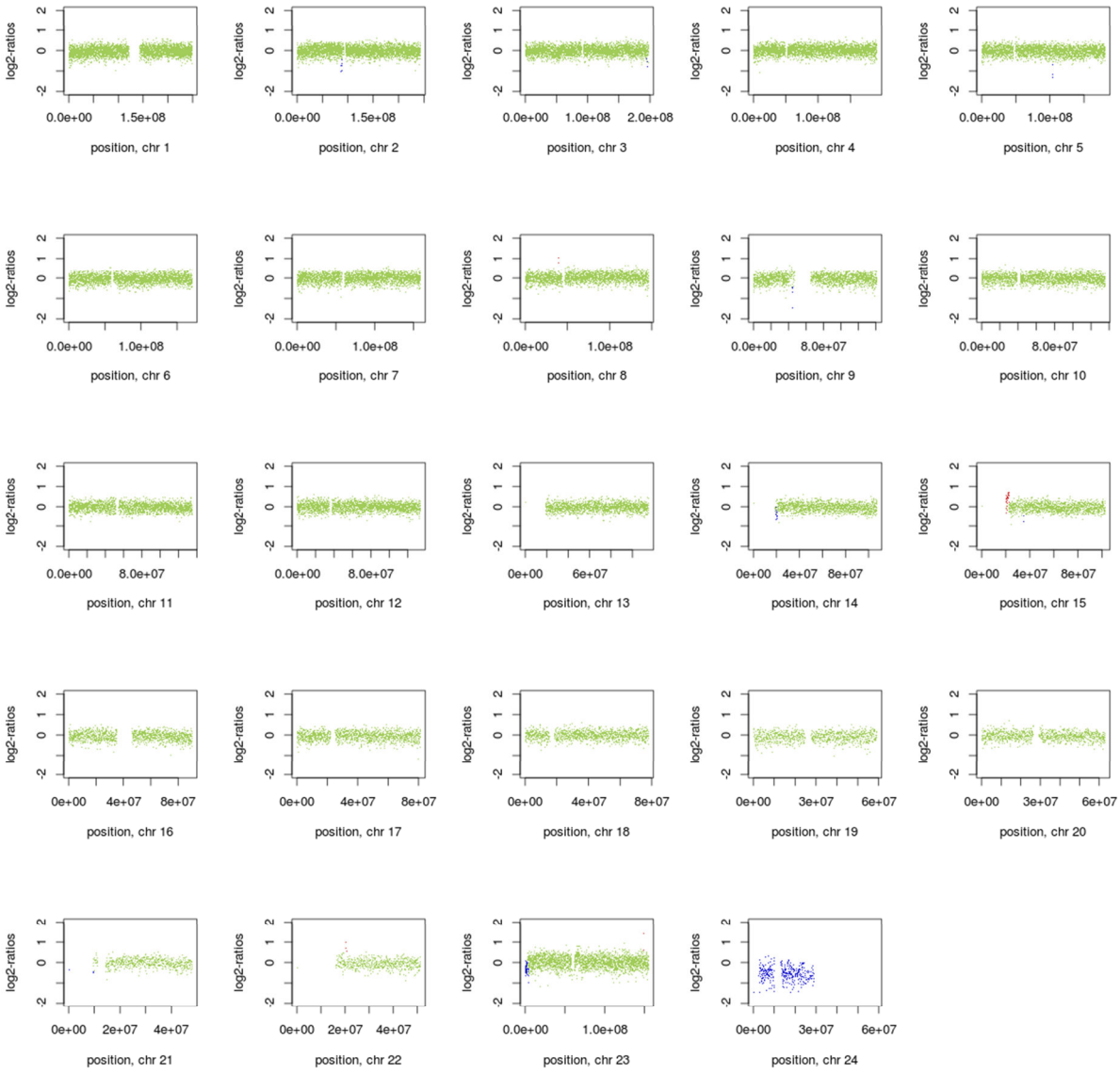
PROSTATE CANCER SAMPLES: PLASMASEQ PROFILES



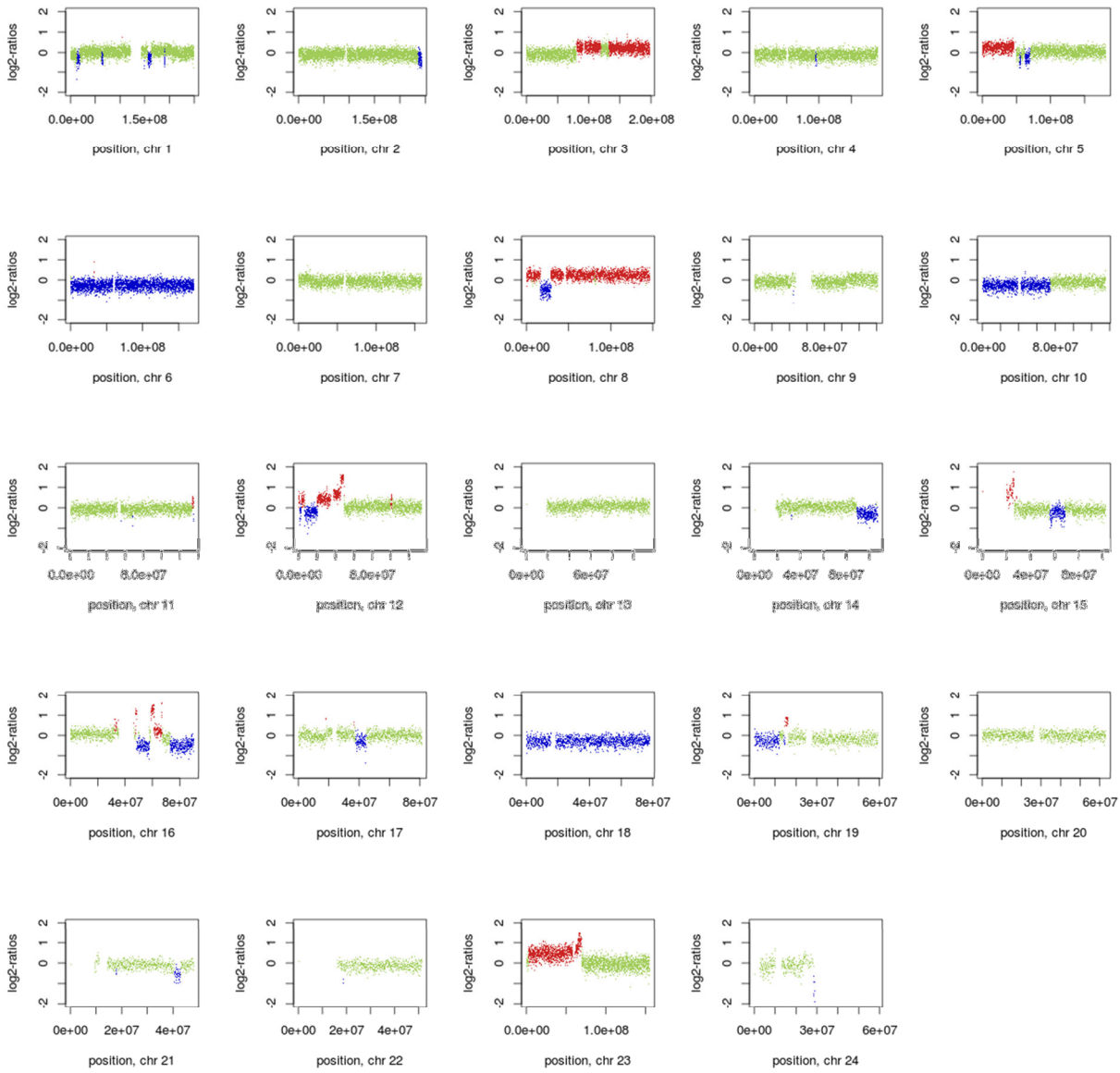
Patient 40u


log2-ratios
position, chr 1
position, chr 2
position, chr 3
position, chr 4
position, chr 5
position, chr 6
position, chr 7
position, chr 8
position, chr 9
position, chr 10
position, chr 11
position, chr 12
position, chr 13
position, chr 14
position, chr 15
position, chr 16
position, chr 17
position, chr 18
position, chr 19
position, chr 20
position, chr 21
position, chr 22
position, chr 23
position, chr 24

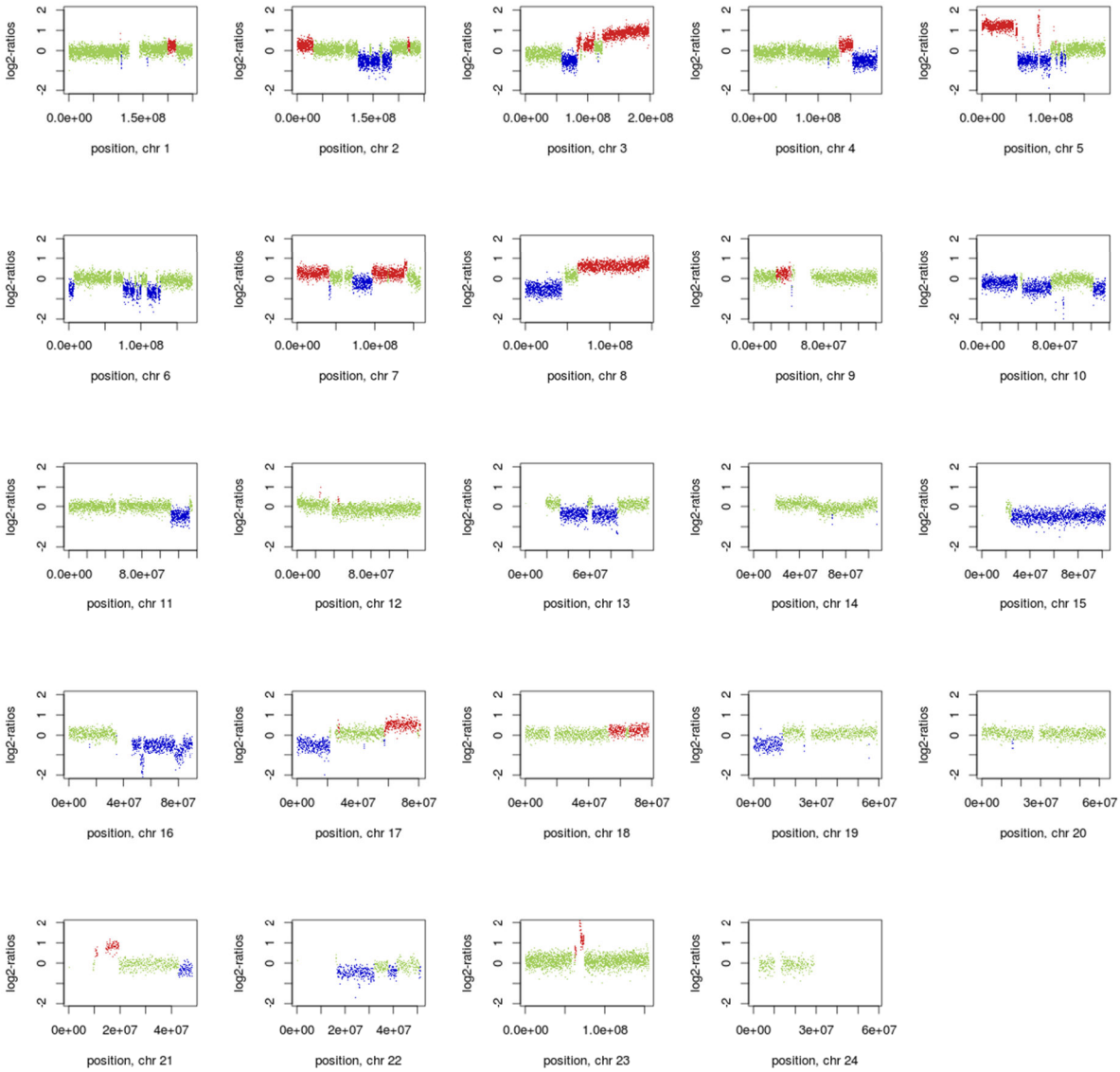
Patient 67


log2-ratios
position, chr 1
position, chr 2
position, chr 3
position, chr 4
position, chr 5
position, chr 6
position, chr 7
position, chr 8
position, chr 9
position, chr 10
position, chr 11
position, chr 12
position, chr 13
position, chr 14
position, chr 15
position, chr 16
position, chr 17
position, chr 18
position, chr 19
position, chr 20
position, chr 21
position, chr 22
position, chr 23
position, chr 24

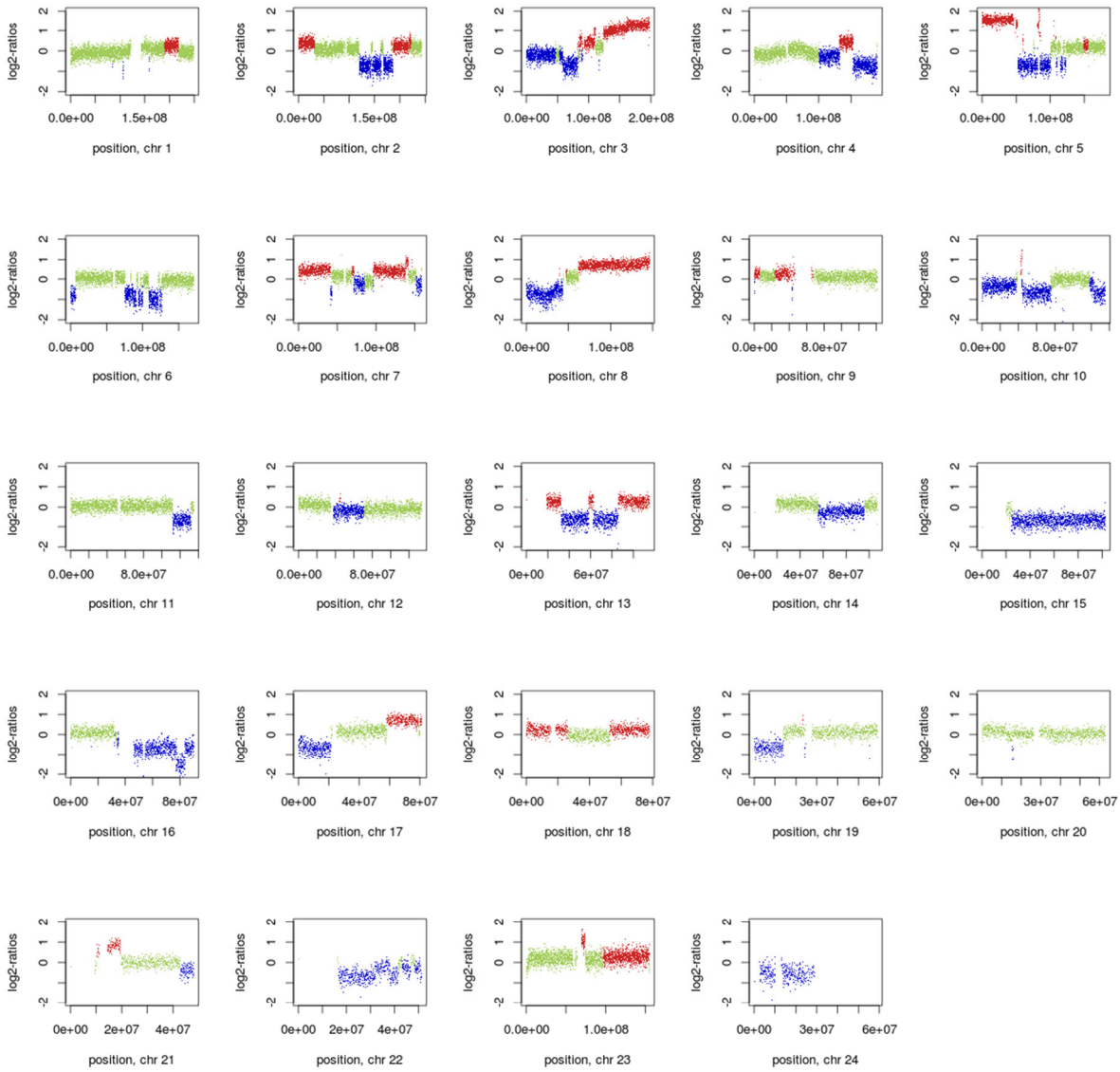
Patient 96



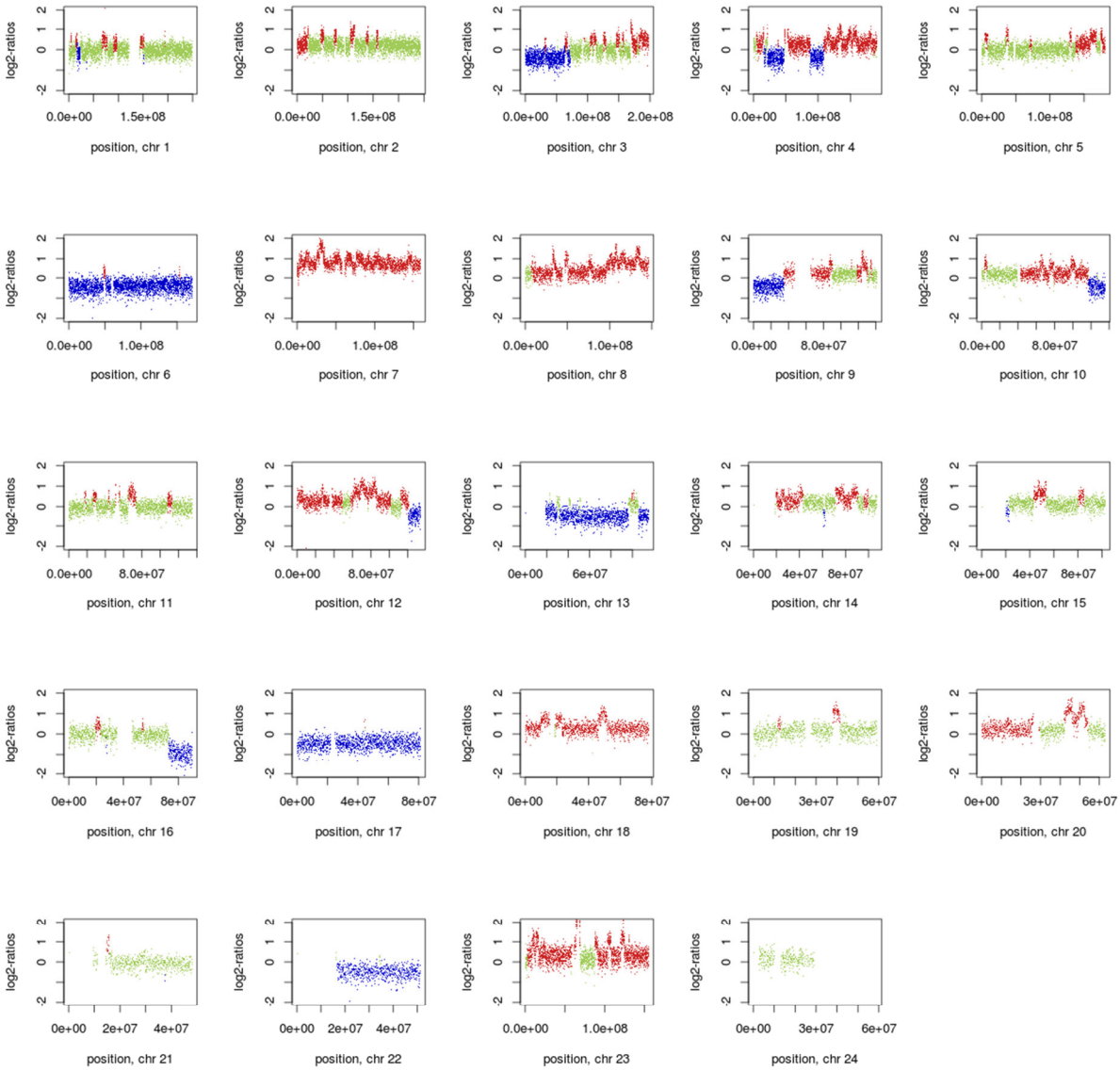
Patient 108u (patient 40u 9 months later)



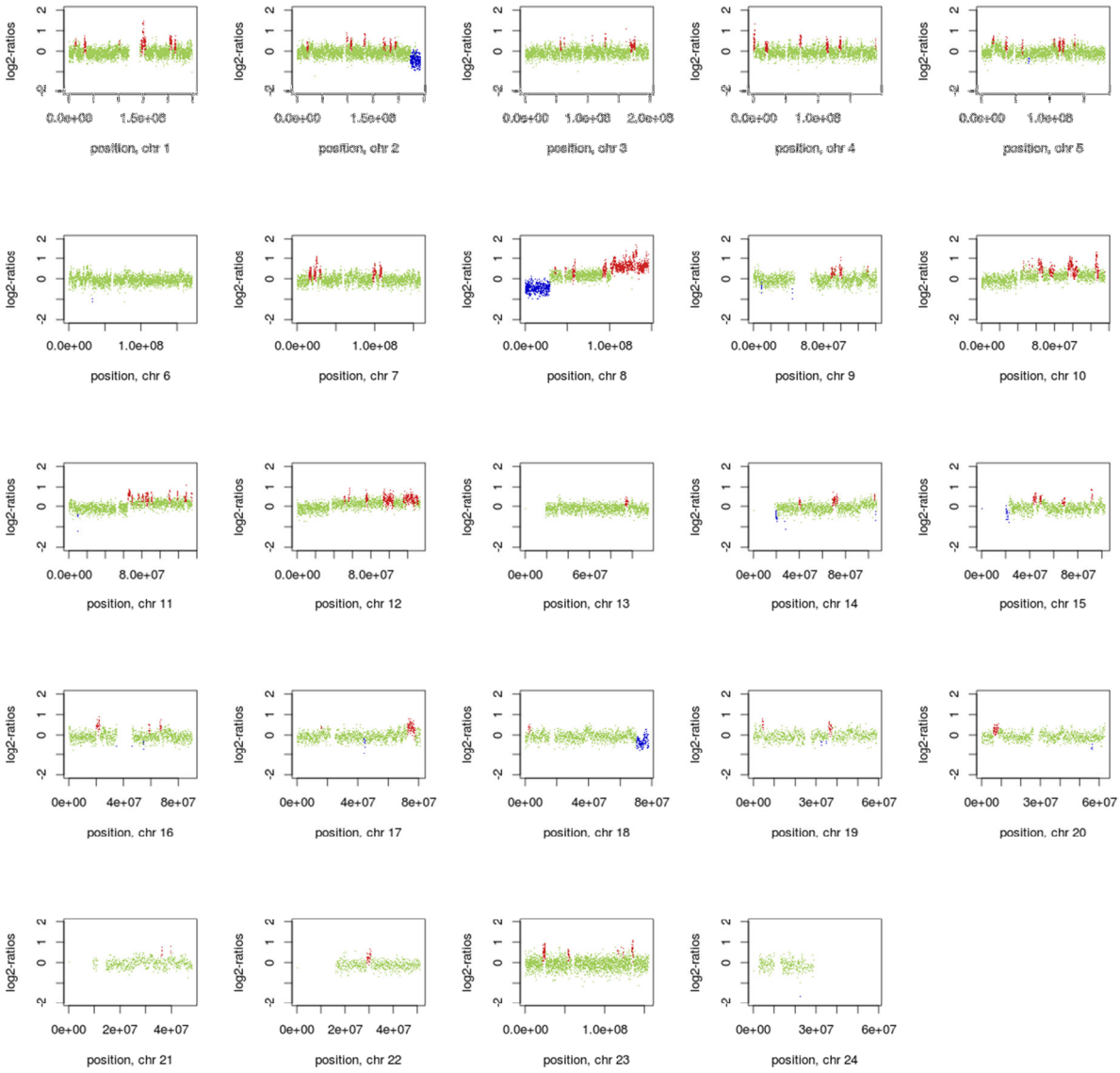
Patient 110o



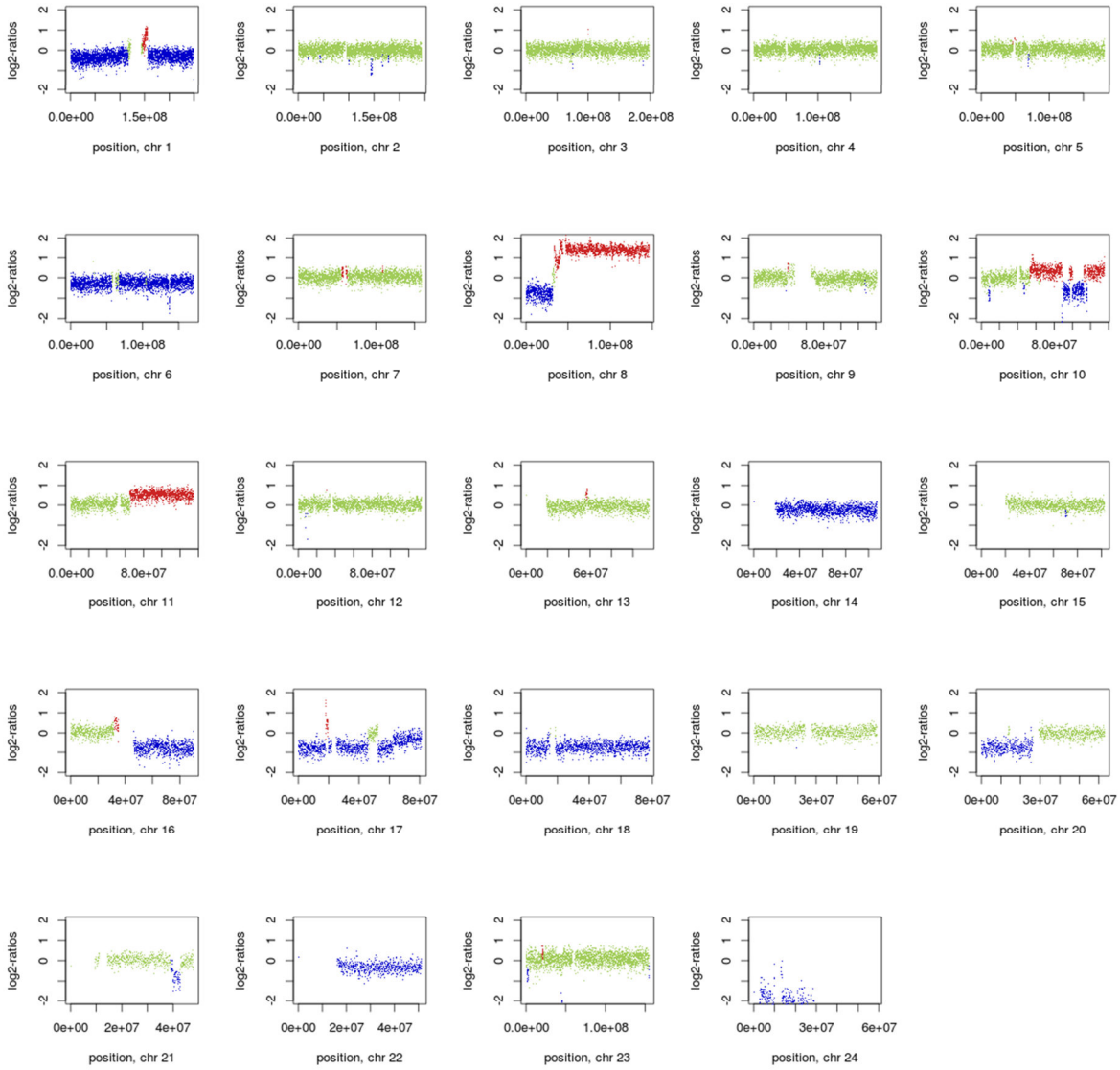
Patient 110_3o (Patient 110o after 6 months)



Patient 117u



Patient 126u



Patient 127u

PROSTATE CANCER SAMPLES: Z-SCORES

Chromosome arm	P8			P9			P15		
	raw	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*
chr1p	1017	0,031	-2,92	2720	0,031	-2,07	4500	0,0323	-0,57
chr1q	1184	0,036	-1,97	3225	0,037	-0,26	5219	0,0375	0,17
chr10p	463	0,014	-1,08	1251	0,014	0,34	1987	0,0143	-0,15
chr10q	982	0,030	-0,78	2665	0,031	1,38	4140	0,0298	-0,72
chr11p	672	0,020	1,48	1740	0,020	0,77	2756	0,0198	0,16
chr11q	909	0,028	0,86	2393	0,028	0,94	3791	0,0272	0,50
chr12p	489	0,015	1,35	1176	0,014	-0,82	1897	0,0136	-0,68
chr12q	1157	0,035	-1,19	3194	0,037	1,61	5002	0,0360	0,26
chr13p	0	0,000	0,00	0	0,000	0,00	0	0,0000	0,00
chr13q	1296	0,039	2,46	3234	0,037	-0,53	5279	0,0379	0,49
chr14p	0	0,000	0,00	0	0,000	0,00	0	0,0000	0,00
chr14q	924	0,028	-1,71	2436	0,028	-1,49	4055	0,0291	0,68
chr15p	0	0,000	0,00	0	0,000	0,00	0	0,0000	0,00
chr15q	772	0,023	3,10	1864	0,021	-0,50	2899	0,0208	-1,74
chr16p	332	0,010	0,57	844	0,010	-0,29	1389	0,0100	0,39
chr16q	352	0,011	-3,85	1096	0,013	0,84	1758	0,0126	0,84
chr17p	189	0,006	0,01	477	0,005	-0,50	826	0,0059	0,51
chr17q	402	0,012	0,67	963	0,011	-1,29	1706	0,0123	0,84
chr18p	167	0,005	0,94	444	0,005	1,28	718	0,0052	1,52
chr18q	703	0,021	-0,83	1934	0,022	0,63	3106	0,0223	0,68
chr19p	120	0,004	0,63	286	0,003	-0,42	469	0,0034	-0,19
chr19q	245	0,007	0,31	633	0,007	0,00	993	0,0071	-0,41
chr2p	951	0,029	-1,12	2460	0,028	-1,73	3974	0,0286	-1,44
chr2q	1895	0,057	0,53	4865	0,056	-1,51	7906	0,0568	-0,33
chr20p	271	0,008	-1,41	696	0,008	-1,92	1165	0,0084	-0,93
chr20q	182	0,006	-1,71	520	0,006	0,04	843	0,0061	0,28
chr21p	46	0,001	-0,16	110	0,001	-0,91	228	0,0016	1,33
chr21q	463	0,014	-1,38	1297	0,015	0,74	2057	0,0148	0,37
chr22p	0	0,000	0,00	0	0,000	0,00	0	0,0000	0,00
chr22q	183	0,006	0,77	486	0,006	0,95	703	0,0051	-0,68
chr3p	985	0,030	-1,78	2669	0,031	-0,29	4252	0,0306	-0,61
chr3q	1489	0,045	-0,62	3959	0,046	0,02	6439	0,0463	0,78
chr4p	658	0,020	-0,38	1762	0,020	0,17	2807	0,0202	-0,02
chr4q	2219	0,067	2,18	5741	0,066	1,36	9062	0,0651	0,50
chr5p	732	0,022	2,15	1785	0,021	-1,15	2961	0,0213	0,32
chr5q	1740	0,053	-0,92	4583	0,053	-0,77	7451	0,0536	-0,05
chr6p	624	0,019	-1,16	1630	0,019	-1,39	2638	0,0190	-1,00
chr6q	1609	0,049	2,07	4154	0,048	0,43	6680	0,0480	0,69
chr7p	641	0,019	1,67	1611	0,019	0,09	2609	0,0188	0,43
chr7q	1343	0,041	5,74	3341	0,039	2,59	5065	0,0364	-0,50
chr8p	492	0,015	0,86	1307	0,015	1,45	2156	0,0155	2,95
chr8q	1393	0,042	0,25	3602	0,042	-0,81	5765	0,0414	-0,95
chr9p	531	0,016	-2,14	1503	0,017	0,89	2326	0,0167	-0,58
chr9q	671	0,020	-0,96	1756	0,020	-1,11	3036	0,0218	2,01
chrXp	397	0,012	-2,95	1165	0,013	0,65	1735	0,0125	-1,80
chrXq	997	0,030	-0,78	2746	0,032	1,03	4187	0,0301	-0,89
chrYp	64	0,002	-1,11	197	0,002	0,50	285	0,0020	-0,57
chrYq	82	0,002	0,86	237	0,003	1,77	310	0,0022	-0,08

Chromosome arm	P16			P18			P23		
	raw	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*
chr1p	3962	0,0322	-0,78	2486	0,0282	-6,89	4151	0,0329	0,27
chr1q	4459	0,0362	-1,46	5531	0,0627	32,51	4625	0,0367	-0,93
chr10p	1806	0,0147	1,24	1195	0,0135	-2,76	1812	0,0144	0,13
chr10q	3668	0,0298	-0,60	2457	0,0278	-4,91	3561	0,0282	-4,07
chr11p	2410	0,0196	-0,39	1493	0,0169	-7,01	2588	0,0205	1,90
chr11q	3261	0,0265	-0,47	1665	0,0189	-10,48	3511	0,0278	1,26
chr12p	1873	0,0152	2,09	1323	0,0150	1,68	1804	0,0143	0,47
chr12q	4326	0,0352	-0,99	3467	0,0393	5,46	4509	0,0357	-0,08
chr13p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr13q	4648	0,0378	0,23	3009	0,0341	-5,40	4793	0,0380	0,56
chr14p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr14q	3561	0,0289	0,26	2397	0,0272	-3,37	3515	0,0279	-1,94
chr15p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr15q	2647	0,0215	-0,45	1697	0,0192	-4,82	2676	0,0212	-1,03
chr16p	1227	0,0100	0,36	1056	0,0120	5,64	1264	0,0100	0,48
chr16q	1592	0,0129	1,56	814	0,0092	-7,25	1699	0,0135	2,81
chr17p	729	0,0059	0,48	599	0,0068	2,47	684	0,0054	-0,68
chr17q	1441	0,0117	-0,17	1089	0,0123	0,98	1566	0,0124	1,11
chr18p	596	0,0048	-0,23	368	0,0042	-3,93	567	0,0045	-2,15
chr18q	2763	0,0225	0,87	1695	0,0192	-3,82	2770	0,0220	0,14
chr19p	413	0,0034	-0,23	337	0,0038	1,21	405	0,0032	-0,69
chr19q	802	0,0065	-1,99	595	0,0067	-1,42	981	0,0078	1,22
chr2p	3555	0,0289	-0,97	2443	0,0277	-2,70	3580	0,0284	-1,71
chr2q	7090	0,0576	0,93	4521	0,0512	-9,16	7441	0,0590	3,06
chr20p	1036	0,0084	-0,80	830	0,0094	1,98	1023	0,0081	-1,68
chr20q	718	0,0058	-0,53	547	0,0062	0,78	746	0,0059	-0,25
chr21p	188	0,0015	0,66	102	0,0012	-1,59	180	0,0014	0,05
chr21q	1836	0,0149	0,68	1142	0,0129	-3,82	1920	0,0152	1,35
chr22p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr22q	703	0,0057	1,28	324	0,0037	-4,79	644	0,0051	-0,53
chr3p	3723	0,0303	-1,09	2434	0,0276	-5,32	3922	0,0311	0,21
chr3q	5594	0,0455	-0,17	3810	0,0432	-2,85	5922	0,0469	1,54
chr4p	2590	0,0210	1,20	1321	0,0150	-7,30	2690	0,0213	1,58
chr4q	8159	0,0663	1,47	4561	0,0517	-10,56	8178	0,0648	0,24
chr5p	2586	0,0210	-0,23	1676	0,0190	-4,45	2689	0,0213	0,39
chr5q	6656	0,0541	0,48	4341	0,0492	-4,38	6682	0,0530	-0,64
chr6p	2348	0,0191	-0,73	2124	0,0241	10,39	2358	0,0187	-1,61
chr6q	5723	0,0465	-2,27	5208	0,0590	22,40	6128	0,0486	1,79
chr7p	2334	0,0190	0,85	1723	0,0195	1,89	2241	0,0178	-1,45
chr7q	4603	0,0374	0,98	3333	0,0378	1,49	4642	0,0368	0,07
chr8p	1847	0,0150	1,26	1137	0,0129	-6,14	1885	0,0149	1,02
chr8q	5108	0,0415	-0,82	3365	0,0381	-6,34	5185	0,0411	-1,50
chr9p	2090	0,0170	0,07	1209	0,0137	-7,91	2124	0,0168	-0,30
chr9q	2594	0,0211	0,55	1426	0,0162	-9,18	2570	0,0204	-0,86
chrXp	1543	0,0125	-1,62	1968	0,0223	23,29	1557	0,0123	-2,13
chrXq	3681	0,0299	-1,11	5400	0,0612	37,44	3845	0,0305	-0,42
chrYp	268	0,0022	0,05	48	0,0005	-7,81	240	0,0019	-1,28
chrYq	286	0,0023	0,28	2	0,0000	-8,18	304	0,0024	0,59

Chromosome arm	P24			P25			P33		
	raw	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*
chr1p	4164	0,0316	-1,73	3217	0,0315	-1,79	3640	0,0322	-0,79
chr1q	4981	0,0378	0,51	3755	0,0368	-0,73	4410	0,0390	2,09
chr10p	1933	0,0147	1,18	1556	0,0153	3,27	1552	0,0137	-2,10
chr10q	3916	0,0297	-0,85	2957	0,0290	-2,39	3426	0,0303	0,47
chr11p	2485	0,0188	-2,23	2033	0,0199	0,46	2279	0,0202	1,02
chr11q	3603	0,0273	0,60	2774	0,0272	0,44	3095	0,0274	0,67
chr12p	1901	0,0144	0,68	1359	0,0133	-1,22	1579	0,0140	-0,10
chr12q	4742	0,0360	0,27	3745	0,0367	1,45	4058	0,0359	0,16
chr13p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr13q	5047	0,0383	1,00	3937	0,0386	1,49	4303	0,0381	0,67
chr14p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr14q	3730	0,0283	-1,07	2978	0,0292	0,78	3132	0,0277	-2,26
chr15p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr15q	2820	0,0214	-0,69	2255	0,0221	0,69	2407	0,0213	-0,88
chr16p	1251	0,0095	-0,92	969	0,0095	-0,89	1092	0,0097	-0,47
chr16q	1585	0,0120	-0,61	1300	0,0127	1,10	1351	0,0119	-0,78
chr17p	702	0,0053	-0,90	542	0,0053	-0,93	614	0,0054	-0,66
chr17q	1643	0,0125	1,20	1159	0,0114	-0,81	1260	0,0111	-1,20
chr18p	686	0,0052	1,74	511	0,0050	0,69	543	0,0048	-0,45
chr18q	2805	0,0213	-0,84	2286	0,0224	0,80	2384	0,0211	-1,11
chr19p	473	0,0036	0,49	329	0,0032	-0,64	346	0,0031	-1,16
chr19q	952	0,0072	-0,20	759	0,0074	0,37	797	0,0070	-0,63
chr2p	3808	0,0289	-0,99	2924	0,0287	-1,29	3153	0,0279	-2,40
chr2q	7584	0,0575	0,76	5803	0,0569	-0,22	6366	0,0563	-1,14
chr20p	1180	0,0089	0,69	876	0,0086	-0,33	993	0,0088	0,22
chr20q	816	0,0062	0,75	609	0,0060	-0,04	625	0,0055	-1,64
chr21p	197	0,0015	0,46	157	0,0015	0,73	165	0,0015	0,25
chr21q	1943	0,0147	0,25	1448	0,0142	-0,97	1574	0,0139	-1,59
chr22p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr22q	693	0,0053	-0,08	515	0,0050	-0,69	568	0,0050	-0,77
chr3p	4086	0,0310	0,06	3131	0,0307	-0,40	3434	0,0304	-0,91
chr3q	5939	0,0450	-0,67	4605	0,0451	-0,54	5233	0,0463	0,78
chr4p	2719	0,0206	0,60	2165	0,0212	1,44	2284	0,0202	0,01
chr4q	8673	0,0658	1,03	6814	0,0668	1,87	7269	0,0643	-0,19
chr5p	2842	0,0216	0,89	2149	0,0211	-0,12	2522	0,0223	2,46
chr5q	6962	0,0528	-0,80	5328	0,0522	-1,36	5897	0,0522	-1,43
chr6p	2516	0,0191	-0,73	1898	0,0186	-1,79	2147	0,0190	-0,94
chr6q	6314	0,0479	0,44	4832	0,0474	-0,57	5225	0,0462	-2,86
chr7p	2433	0,0185	-0,14	1855	0,0182	-0,64	1967	0,0174	-2,14
chr7q	4858	0,0368	0,14	3824	0,0375	1,09	4319	0,0382	2,14
chr8p	2044	0,0155	2,97	1556	0,0153	2,11	1643	0,0145	-0,40
chr8q	5527	0,0419	-0,17	4357	0,0427	1,13	5080	0,0449	4,74
chr9p	2277	0,0173	0,76	1689	0,0166	-0,97	1957	0,0173	0,86
chr9q	2775	0,0210	0,48	2129	0,0209	0,14	2414	0,0214	1,09
chrXp	1649	0,0125	-1,71	1318	0,0129	-0,65	1615	0,0143	2,83
chrXq	4031	0,0306	-0,30	3130	0,0307	-0,16	3807	0,0337	3,52
chrYp	253	0,0019	-1,20	210	0,0021	-0,52	257	0,0023	0,51
chrYq	327	0,0025	0,85	260	0,0025	1,10	279	0,0025	0,80

Chromosome arm	P36			P37			P39		
	raw	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*
chr1p	4164	0,0314	-1,95	2273	0,0321	-0,93	918	0,0317	-1,50
chr1q	4805	0,0363	-1,43	2547	0,0360	-1,80	1020	0,0353	-2,73
chr10p	1856	0,0140	-1,11	1021	0,0144	0,34	433	0,0150	2,25
chr10q	3907	0,0295	-1,31	2112	0,0298	-0,56	816	0,0282	-4,11
chr11p	2609	0,0197	-0,13	1399	0,0198	0,04	595	0,0206	2,03
chr11q	3505	0,0265	-0,54	1924	0,0272	0,41	750	0,0259	-1,24
chr12p	1838	0,0139	-0,27	1019	0,0144	0,64	436	0,0151	1,82
chr12q	4428	0,0334	-3,71	2579	0,0364	1,00	1007	0,0348	-1,54
chr13p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr13q	4915	0,0371	-0,80	2591	0,0366	-1,57	1088	0,0376	-0,03
chr14p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr14q	3719	0,0281	-1,51	2094	0,0296	1,55	836	0,0289	0,16
chr15p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr15q	2735	0,0206	-2,12	1495	0,0211	-1,21	616	0,0213	-0,88
chr16p	1088	0,0082	-4,30	704	0,0099	0,28	280	0,0097	-0,42
chr16q	1562	0,0118	-1,16	895	0,0126	0,86	371	0,0128	1,29
chr17p	655	0,0049	-1,78	389	0,0055	-0,51	168	0,0058	0,21
chr17q	1390	0,0105	-2,40	842	0,0119	0,16	315	0,0109	-1,68
chr18p	590	0,0045	-2,37	324	0,0046	-1,70	128	0,0044	-2,54
chr18q	2947	0,0222	0,56	1522	0,0215	-0,51	642	0,0222	0,48
chr19p	416	0,0031	-0,91	229	0,0032	-0,62	101	0,0035	0,19
chr19q	885	0,0067	-1,58	549	0,0078	1,17	194	0,0067	-1,51
chr2p	3783	0,0286	-1,45	2060	0,0291	-0,68	861	0,0298	0,26
chr2q	7426	0,0560	-1,55	4091	0,0578	1,18	1691	0,0584	2,22
chr20p	1082	0,0082	-1,52	555	0,0078	-2,44	252	0,0087	0,02
chr20q	648	0,0049	-3,95	379	0,0054	-2,28	163	0,0056	-1,26
chr21p	202	0,0015	0,64	110	0,0016	0,82	41	0,0014	-0,01
chr21q	2003	0,0151	1,12	1078	0,0152	1,37	408	0,0141	-1,19
chr22p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr22q	544	0,0041	-3,49	326	0,0046	-2,01	126	0,0044	-2,76
chr3p	3940	0,0297	-1,91	2143	0,0303	-1,07	922	0,0319	1,44
chr3q	5901	0,0445	-1,25	3271	0,0462	0,68	1340	0,0463	0,81
chr4p	2677	0,0202	0,02	1408	0,0199	-0,43	568	0,0196	-0,78
chr4q	8734	0,0659	1,15	4872	0,0688	3,52	1963	0,0678	2,72
chr5p	2842	0,0214	0,67	1535	0,0217	1,15	641	0,0222	2,14
chr5q	6893	0,0520	-1,57	3726	0,0526	-0,97	1564	0,0541	0,44
chr6p	2529	0,0191	-0,72	1362	0,0192	-0,38	555	0,0192	-0,51
chr6q	6507	0,0491	2,86	3334	0,0471	-1,13	1346	0,0465	-2,26
chr7p	2351	0,0177	-1,48	1309	0,0185	-0,07	515	0,0178	-1,38
chr7q	4740	0,0358	-1,42	2656	0,0375	1,13	1060	0,0366	-0,16
chr8p	1820	0,0137	-3,17	1070	0,0151	1,62	460	0,0159	4,35
chr8q	5245	0,0396	-3,96	3069	0,0433	2,16	1229	0,0425	0,74
chr9p	2265	0,0171	0,34	1221	0,0172	0,70	510	0,0176	1,62
chr9q	2594	0,0196	-2,42	1426	0,0201	-1,30	619	0,0214	1,17
chrXp	2768	0,0209	19,70	857	0,0121	-2,73	364	0,0126	-1,52
chrXq	6701	0,0506	24,37	2145	0,0303	-0,64	883	0,0305	-0,37
chrYp	164	0,0012	-4,47	133	0,0019	-1,39	72	0,0025	1,54
chrYq	124	0,0009	-4,82	157	0,0022	-0,11	69	0,0024	0,50

Chromosome arm	P64			P70			P71		
	raw	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*
chr1p	2213	0,0312	-2,32	2359	0,0316	-1,72	2866	0,0321	-0,96
chr1q	2527	0,0356	-2,27	2757	0,0369	-0,60	3351	0,0375	0,18
chr10p	1058	0,0149	2,06	1112	0,0149	1,97	1374	0,0154	3,72
chr10q	2134	0,0301	-0,03	2168	0,0290	-2,32	2593	0,0290	-2,30
chr11p	1442	0,0203	1,43	1489	0,0199	0,47	1757	0,0197	-0,18
chr11q	1946	0,0274	0,74	2004	0,0268	-0,05	2450	0,0274	0,74
chr12p	961	0,0135	-0,84	1077	0,0144	0,68	1230	0,0138	-0,44
chr12q	2542	0,0358	0,06	2697	0,0361	0,49	3153	0,0353	-0,76
chr13p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr13q	2610	0,0368	-1,28	2757	0,0369	-1,09	3430	0,0384	1,19
chr14p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr14q	1932	0,0272	-3,22	2214	0,0296	1,68	2577	0,0289	0,08
chr15p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr15q	1490	0,0210	-1,43	1679	0,0225	1,39	1964	0,0220	0,46
chr16p	648	0,0091	-1,86	689	0,0092	-1,62	887	0,0099	0,25
chr16q	804	0,0113	-2,25	930	0,0124	0,40	1111	0,0124	0,38
chr17p	440	0,0062	1,12	436	0,0058	0,28	508	0,0057	-0,07
chr17q	784	0,0110	-1,38	840	0,0112	-1,02	1119	0,0125	1,33
chr18p	346	0,0049	-0,05	396	0,0053	2,29	464	0,0052	1,70
chr18q	1499	0,0211	-1,05	1623	0,0217	-0,18	1975	0,0221	0,37
chr19p	239	0,0034	-0,20	247	0,0033	-0,39	332	0,0037	0,89
chr19q	555	0,0078	1,34	501	0,0067	-1,51	688	0,0077	1,04
chr2p	2083	0,0294	-0,31	2230	0,0299	0,39	2613	0,0293	-0,46
chr2q	3998	0,0563	-1,09	4135	0,0554	-2,64	5120	0,0573	0,46
chr20p	608	0,0086	-0,38	629	0,0084	-0,80	776	0,0087	-0,04
chr20q	430	0,0061	0,28	446	0,0060	-0,04	502	0,0056	-1,31
chr21p	85	0,0012	-1,33	101	0,0014	-0,40	120	0,0013	-0,45
chr21q	1124	0,0158	2,76	1064	0,0142	-0,86	1341	0,0150	0,89
chr22p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr22q	346	0,0049	-1,21	398	0,0053	0,14	490	0,0055	0,61
chr3p	2097	0,0296	-2,20	2255	0,0302	-1,20	2764	0,0309	-0,01
chr3q	3263	0,0460	0,43	3274	0,0438	-2,07	3921	0,0439	-1,99
chr4p	1435	0,0202	0,05	1517	0,0203	0,16	1776	0,0199	-0,43
chr4q	4598	0,0648	0,23	4930	0,0660	1,21	5685	0,0636	-0,72
chr5p	1535	0,0216	1,05	1551	0,0208	-0,76	1936	0,0217	1,14
chr5q	3779	0,0533	-0,35	3924	0,0525	-1,07	4676	0,0524	-1,24
chr6p	1377	0,0194	-0,01	1451	0,0194	0,03	1746	0,0195	0,31
chr6q	3448	0,0486	1,84	3517	0,0471	-1,14	4229	0,0473	-0,62
chr7p	1286	0,0181	-0,76	1374	0,0184	-0,25	1568	0,0176	-1,84
chr7q	3304	0,0466	14,42	3271	0,0438	10,34	3749	0,0420	7,68
chr8p	1062	0,0150	1,11	1103	0,0148	0,41	1259	0,0141	-1,92
chr8q	3000	0,0423	0,42	3167	0,0424	0,62	3702	0,0414	-0,93
chr9p	1204	0,0170	0,03	1264	0,0169	-0,09	1519	0,0170	0,12
chr9q	1417	0,0200	-1,64	1520	0,0203	-0,90	1841	0,0206	-0,38
chrXp	899	0,0127	-1,29	967	0,0129	-0,59	1112	0,0124	-1,85
chrXq	2109	0,0297	-1,35	2287	0,0306	-0,25	2691	0,0301	-0,85
chrYp	141	0,0020	-0,87	150	0,0020	-0,77	181	0,0020	-0,68
chrYq	160	0,0023	0,02	200	0,0027	1,57	174	0,0019	-1,10

Chromosome arm	P81			P83			P86		
	raw	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*
chr1p	984	0,0323	-0,64	3952	0,0325	-0,38	1229	0,0289	-5,71
chr1q	1133	0,0372	-0,24	4657	0,0383	1,13	1428	0,0336	-4,82
chr10p	404	0,0133	-3,73	1778	0,0146	0,99	533	0,0126	-6,22
chr10q	925	0,0304	0,60	3483	0,0286	-3,21	1179	0,0278	-5,06
chr11p	608	0,0200	0,53	2425	0,0199	0,44	778	0,0183	-3,52
chr11q	862	0,0283	1,87	3226	0,0265	-0,47	1019	0,0240	-3,76
chr12p	411	0,0135	-0,93	1644	0,0135	-0,91	3829	0,0902	132,70
chr12q	1029	0,0338	-3,15	4574	0,0376	2,80	1456	0,0343	-2,35
chr13p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr13q	1184	0,0389	1,89	4283	0,0352	-3,72	1503	0,0354	-3,40
chr14p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr14q	925	0,0304	3,15	3479	0,0286	-0,47	1148	0,0270	-3,62
chr15p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr15q	696	0,0228	2,10	2536	0,0208	-1,75	833	0,0196	-4,08
chr16p	295	0,0097	-0,41	1096	0,0090	-2,20	365	0,0086	-3,29
chr16q	361	0,0118	-1,02	1455	0,0120	-0,77	475	0,0112	-2,59
chr17p	202	0,0066	2,11	692	0,0057	-0,07	222	0,0052	-1,13
chr17q	374	0,0123	0,86	1481	0,0122	0,67	453	0,0107	-2,08
chr18p	156	0,0051	1,29	607	0,0050	0,56	192	0,0045	-2,00
chr18q	626	0,0205	-1,89	2608	0,0214	-0,62	913	0,0215	-0,51
chr19p	90	0,0030	-1,49	418	0,0034	0,01	130	0,0031	-1,16
chr19q	216	0,0071	-0,53	869	0,0071	-0,40	294	0,0069	-0,95
chr2p	921	0,0302	0,93	3733	0,0307	1,55	1112	0,0262	-4,81
chr2q	1714	0,0563	-1,22	7268	0,0597	4,22	2284	0,0538	-5,12
chr20p	278	0,0091	1,19	1044	0,0086	-0,36	322	0,0076	-3,17
chr20q	188	0,0062	0,68	766	0,0063	1,13	240	0,0057	-1,20
chr21p	59	0,0019	3,13	170	0,0014	-0,13	51	0,0012	-1,31
chr21q	425	0,0139	-1,53	1748	0,0144	-0,60	585	0,0138	-1,92
chr22p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr22q	143	0,0047	-1,75	614	0,0050	-0,71	218	0,0051	-0,44
chr3p	970	0,0318	1,40	4083	0,0335	4,08	1163	0,0274	-5,61
chr3q	1305	0,0428	-3,23	5686	0,0467	1,28	1864	0,0439	-2,00
chr4p	603	0,0198	-0,56	2481	0,0204	0,27	840	0,0198	-0,57
chr4q	1998	0,0656	0,87	7868	0,0646	0,09	2573	0,0606	-3,23
chr5p	683	0,0224	2,69	2500	0,0205	-1,22	848	0,0200	-2,41
chr5q	1648	0,0541	0,48	6472	0,0532	-0,43	2058	0,0485	-5,09
chr6p	597	0,0196	0,41	2334	0,0192	-0,52	762	0,0179	-3,27
chr6q	1480	0,0486	1,81	5718	0,0470	-1,36	1828	0,0430	-9,11
chr7p	591	0,0194	1,66	2216	0,0182	-0,60	712	0,0168	-3,33
chr7q	1158	0,0380	1,86	4378	0,0360	-1,14	1467	0,0345	-3,23
chr8p	446	0,0146	-0,03	1749	0,0144	-0,97	608	0,0143	-1,15
chr8q	1243	0,0408	-1,98	5283	0,0434	2,25	1688	0,0398	-3,69
chr9p	478	0,0157	-3,07	2020	0,0166	-0,88	664	0,0156	-3,20
chr9q	576	0,0189	-3,74	2506	0,0206	-0,42	794	0,0187	-4,15
chrXp	397	0,0130	-0,37	1546	0,0127	-1,21	468	0,0110	-5,50
chrXq	928	0,0305	-0,44	3702	0,0304	-0,50	1156	0,0272	-4,43
chrYp	82	0,0027	2,52	270	0,0022	0,24	78	0,0018	-1,59
chrYq	75	0,0025	0,78	307	0,0025	1,00	103	0,0024	0,65

Chromosome arm	P96			P40u			P67		
	raw	ratio	Z-score	raw	ratio	Z-score	raw	ratio	Z-score
chr1p	2336	0,0311	-2,47	8472	0,0339	1,76	12324	0,0293	-5,13
chr1q	2757	0,0367	-0,88	9907	0,0396	2,89	13298	0,0316	-7,37
chr10p	1014	0,0135	-2,91	2857	0,0114	-10,17	6381	0,0152	3,02
chr10q	2121	0,0282	-4,05	6447	0,0258	-9,38	13068	0,0311	2,20
chr11p	1481	0,0197	-0,08	4630	0,0185	-3,04	7254	0,0173	-6,15
chr11q	2024	0,0269	0,10	6371	0,0255	-1,81	9848	0,0234	-4,49
chr12p	1030	0,0137	-0,55	3900	0,0156	2,74	5920	0,0141	0,11
chr12q	2607	0,0347	-1,71	10966	0,0439	12,63	17193	0,0409	8,02
chr13p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr13q	2801	0,0373	-0,53	9814	0,0393	2,49	10483	0,0249	-19,36
chr14p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr14q	2163	0,0288	-0,05	7576	0,0303	3,03	12903	0,0307	3,85
chr15p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr15q	1639	0,0218	0,12	5277	0,0211	-1,23	9352	0,0223	0,97
chr16p	713	0,0095	-0,92	2701	0,0108	2,56	3567	0,0085	-3,57
chr16q	857	0,0114	-2,07	2989	0,0120	-0,77	3738	0,0089	-8,02
chr17p	382	0,0051	-1,46	1557	0,0062	1,18	1582	0,0038	-4,50
chr17q	897	0,0119	0,25	2900	0,0116	-0,37	3231	0,0077	-7,52
chr18p	397	0,0053	2,19	984	0,0039	-5,22	2489	0,0059	5,71
chr18q	1732	0,0231	1,73	4342	0,0174	-6,47	10115	0,0241	3,20
chr19p	248	0,0033	-0,41	1012	0,0040	1,93	1520	0,0036	0,58
chr19q	484	0,0064	-2,19	1689	0,0068	-1,38	3455	0,0082	2,36
chr2p	2146	0,0286	-1,44	6919	0,0277	-2,70	13146	0,0313	2,42
chr2q	4261	0,0567	-0,51	13306	0,0532	-6,01	24550	0,0584	2,19
chr20p	656	0,0087	0,08	2235	0,0089	0,67	3760	0,0089	0,69
chr20q	434	0,0058	-0,75	1586	0,0063	1,31	3215	0,0077	6,05
chr21p	107	0,0014	0,03	410	0,0016	1,34	716	0,0017	1,72
chr21q	1159	0,0154	1,82	3302	0,0132	-3,21	5124	0,0122	-5,51
chr22p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr22q	409	0,0054	0,48	1186	0,0047	-1,60	1427	0,0034	-5,60
chr3p	2339	0,0311	0,28	7571	0,0303	-1,05	9443	0,0225	-13,36
chr3q	3557	0,0473	2,01	13953	0,0558	11,86	18467	0,0439	-1,94
chr4p	1576	0,0210	1,09	4514	0,0181	-2,99	7200	0,0171	-4,27
chr4q	4975	0,0662	1,38	14840	0,0594	-4,25	28075	0,0668	1,88
chr5p	1561	0,0208	-0,73	6147	0,0246	7,20	8091	0,0193	-3,90
chr5q	3976	0,0529	-0,68	14328	0,0573	3,66	21215	0,0505	-3,09
chr6p	1486	0,0198	0,82	3758	0,0150	-9,77	5495	0,0131	-14,13
chr6q	3551	0,0473	-0,79	9405	0,0376	-19,83	13568	0,0323	-30,36
chr7p	1431	0,0190	0,99	4455	0,0178	-1,34	12480	0,0297	21,21
chr7q	2851	0,0379	1,76	8734	0,0349	-2,66	23210	0,0552	27,14
chr8p	1126	0,0150	1,18	3677	0,0147	0,21	6548	0,0156	3,25
chr8q	3357	0,0447	4,33	13029	0,0521	16,42	23157	0,0551	21,29
chr9p	1297	0,0173	0,74	3914	0,0157	-3,16	4801	0,0114	-13,42
chr9q	1539	0,0205	-0,63	5127	0,0205	-0,58	9081	0,0216	1,60
chrXp	990	0,0132	0,00	3618	0,0145	3,31	6277	0,0149	4,50
chrXq	2440	0,0325	2,04	8528	0,0341	4,06	21855	0,0520	26,13
chrYp	121	0,0016	-2,68	555	0,0022	0,25	755	0,0018	-1,78
chrYq	112	0,0015	-2,78	535	0,0021	-0,40	861	0,0020	-0,73

Chromosome arm	P108u			P110o			P110-3o		
	raw	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*
chr1p	8765	0,0327	0,02	6411	0,0297	-4,49	8723	0,0280	-7,21
chr1q	10217	0,0382	1,00	7748	0,0360	-1,83	11236	0,0360	-1,76
chr10p	3203	0,0120	-8,29	2650	0,0123	-7,12	3251	0,0104	-13,71
chr10q	7377	0,0276	-5,53	5083	0,0236	-14,17	6855	0,0220	-17,70
chr11p	5194	0,0194	-0,86	4017	0,0186	-2,74	5868	0,0188	-2,33
chr11q	7159	0,0267	-0,17	5362	0,0249	-2,60	7364	0,0236	-4,28
chr12p	4416	0,0165	4,30	3171	0,0147	1,20	4265	0,0137	-0,62
chr12q	11044	0,0412	8,54	6700	0,0311	-7,35	9633	0,0309	-7,70
chr13p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr13q	10542	0,0394	2,67	7072	0,0328	-7,34	9315	0,0299	-11,87
chr14p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr14q	7920	0,0296	1,56	6140	0,0285	-0,65	8410	0,0270	-3,79
chr15p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr15q	5693	0,0213	-0,93	3349	0,0155	-11,87	3904	0,0125	-17,66
chr16p	2733	0,0102	0,98	2021	0,0094	-1,21	2795	0,0090	-2,33
chr16q	3255	0,0122	-0,29	1674	0,0078	-10,69	1903	0,0061	-14,65
chr17p	1534	0,0057	0,03	940	0,0044	-3,12	1026	0,0033	-5,60
chr17q	3228	0,0121	0,46	2850	0,0132	2,60	4321	0,0138	3,74
chr18p	1160	0,0043	-3,04	1165	0,0054	2,86	1706	0,0055	3,20
chr18q	4929	0,0184	-4,97	4617	0,0214	-0,62	6549	0,0210	-1,25
chr19p	1063	0,0040	1,69	688	0,0032	-0,75	891	0,0029	-1,80
chr19q	1769	0,0066	-1,76	1598	0,0074	0,30	2130	0,0068	-1,20
chr2p	7436	0,0278	-2,56	6390	0,0297	0,11	9747	0,0312	2,36
chr2q	14537	0,0543	-4,32	10662	0,0495	-11,91	14443	0,0463	-16,94
chr20p	2244	0,0084	-0,91	1824	0,0085	-0,68	2539	0,0081	-1,60
chr20q	1677	0,0063	1,02	1270	0,0059	-0,32	1871	0,0060	0,05
chr21p	457	0,0017	1,74	355	0,0016	1,38	584	0,0019	2,74
chr21q	3536	0,0132	-3,21	3275	0,0152	1,30	4709	0,0151	1,06
chr22p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr22q	1252	0,0047	-1,80	889	0,0041	-3,44	1071	0,0034	-5,50
chr3p	8203	0,0306	-0,50	5671	0,0263	-7,30	7325	0,0235	-11,78
chr3q	13910	0,0519	7,37	15014	0,0697	27,98	25431	0,0815	41,75
chr4p	4743	0,0177	-3,46	3815	0,0177	-3,48	5088	0,0163	-5,43
chr4q	16078	0,0600	-3,68	12122	0,0562	-6,80	16307	0,0523	-10,08
chr5p	6554	0,0245	6,97	9676	0,0449	49,46	17175	0,0550	70,56
chr5q	14670	0,0548	1,17	10593	0,0492	-4,41	15240	0,0488	-4,72
chr6p	4406	0,0165	-6,59	3798	0,0176	-3,98	5510	0,0177	-3,91
chr6q	10945	0,0409	-13,40	8288	0,0385	-18,17	11211	0,0359	-23,17
chr7p	4966	0,0185	0,04	4421	0,0205	3,78	6593	0,0211	4,94
chr7q	9583	0,0358	-1,40	7857	0,0365	-0,42	11912	0,0382	2,10
chr8p	4008	0,0150	1,12	2158	0,0100	-16,11	2474	0,0079	-23,36
chr8q	13303	0,0497	12,47	12579	0,0584	26,60	18767	0,0601	29,48
chr9p	4314	0,0161	-2,05	3769	0,0175	1,29	5504	0,0176	1,65
chr9q	5446	0,0203	-0,91	4459	0,0207	-0,22	6292	0,0202	-1,26
chrXp	4690	0,0175	11,08	3021	0,0140	2,15	4181	0,0134	0,57
chrXq	8611	0,0322	1,66	9545	0,0443	16,62	17110	0,0548	29,62
chrYp	491	0,0018	-1,60	404	0,0019	-1,41	385	0,0012	-4,49
chrYq	505	0,0019	-1,33	399	0,0019	-1,46	439	0,0014	-3,09

Chromosome arms	P117u			P126u			P127u		
	raw	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*	*raw*	*ratio*	*Z-score*
chr1p	1760	0,0309	-2,75	6318	0,0310	-2,54	8760	0,0255	-10,96
chr1q	2208	0,0388	1,79	7611	0,0374	0,02	10533	0,0306	-8,68
chr10p	588	0,0103	-14,04	2750	0,0135	-2,86	4543	0,0132	-3,91
chr10q	1603	0,0281	-4,23	7503	0,0369	14,77	10698	0,0311	2,23
chr11p	1197	0,0210	3,15	3646	0,0179	-4,54	6961	0,0202	1,23
chr11q	1489	0,0261	-0,94	6392	0,0314	5,95	12884	0,0375	13,89
chr12p	835	0,0147	1,11	2578	0,0127	-2,37	4850	0,0141	0,14
chr12q	2006	0,0352	-0,89	8807	0,0433	11,71	12214	0,0355	-0,42
chr13p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr13q	2105	0,0370	-1,01	7118	0,0350	-4,05	11808	0,0343	-5,02
chr14p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr14q	1432	0,0251	-7,46	5764	0,0283	-1,00	8297	0,0241	-9,53
chr15p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr15q	1019	0,0179	-7,37	4320	0,0212	-1,00	7234	0,0210	-1,36
chr16p	541	0,0095	-0,89	1943	0,0095	-0,77	3315	0,0096	-0,52
chr16q	494	0,0087	-8,55	2303	0,0113	-2,29	2432	0,0071	-12,34
chr17p	252	0,0044	-2,98	1094	0,0054	-0,79	1102	0,0032	-5,79
chr17q	695	0,0122	0,73	2174	0,0107	-2,05	2813	0,0082	-6,62
chr18p	274	0,0048	-0,41	998	0,0049	0,10	1069	0,0031	-9,76
chr18q	1312	0,0230	1,71	4122	0,0203	-2,31	4369	0,0127	-13,20
chr19p	212	0,0037	0,91	655	0,0032	-0,67	1147	0,0033	-0,30
chr19q	384	0,0067	-1,42	1456	0,0072	-0,37	2450	0,0071	-0,44
chr2p	1662	0,0292	-0,56	5781	0,0284	-1,67	9640	0,0280	-2,19
chr2q	3500	0,0615	6,97	11103	0,0545	-3,91	19055	0,0554	-2,55
chr20p	471	0,0083	-1,22	1750	0,0086	-0,30	1731	0,0050	-10,37
chr20q	365	0,0064	1,55	1114	0,0055	-1,84	2039	0,0059	-0,19
chr21p	73	0,0013	-0,83	281	0,0014	-0,23	431	0,0013	-1,00
chr21q	624	0,0110	-8,31	2910	0,0143	-0,74	4696	0,0137	-2,19
chr22p	0	0,0000	0,00	0	0,0000	0,00	0	0,0000	0,00
chr22q	259	0,0045	-2,18	961	0,0047	-1,67	1338	0,0039	-4,13
chr3p	1750	0,0307	-0,35	5954	0,0293	-2,68	10877	0,0316	1,07
chr3q	2701	0,0474	2,11	9219	0,0453	-0,37	15237	0,0443	-1,52
chr4p	1036	0,0182	-2,80	3954	0,0194	-1,07	6507	0,0189	-1,77
chr4q	3501	0,0615	-2,51	12863	0,0632	-1,09	21992	0,0640	-0,47
chr5p	1109	0,0195	-3,44	4298	0,0211	-0,02	7433	0,0216	1,02
chr5q	2986	0,0524	-1,17	10490	0,0515	-2,05	18681	0,0543	0,71
chr6p	978	0,0172	-4,99	3845	0,0189	-1,16	5213	0,0152	-9,48
chr6q	2215	0,0389	-17,32	9366	0,0460	-3,25	13583	0,0395	-16,12
chr7p	1058	0,0186	0,10	3640	0,0179	-1,21	6181	0,0180	-1,04
chr7q	2233	0,0392	3,62	7196	0,0354	-2,04	12620	0,0367	-0,07
chr8p	668	0,0117	-10,15	2353	0,0116	-10,73	4140	0,0120	-9,07
chr8q	3302	0,0580	25,96	12067	0,0593	28,09	36408	0,1059	103,87
chr9p	826	0,0145	-5,95	3366	0,0165	-1,02	5827	0,0169	-0,03
chr9q	1192	0,0209	0,25	4024	0,0198	-2,04	6802	0,0198	-2,02
chrXp	1053	0,0185	13,57	2679	0,0132	-0,03	4472	0,0130	-0,44
chrXq	2722	0,0478	20,94	6021	0,0296	-1,52	11085	0,0322	1,75
chrYp	108	0,0019	-1,30	381	0,0019	-1,42	274	0,0008	-6,59
chrYq	155	0,0027	1,74	375	0,0018	-1,49	141	0,0004	-6,75

COLON CANCER SAMPLE: Z-SCORES

Chromosome arm	C38		
	raw	*ratio*	*Z-score*
chr1p	4983	0,03112877	-2,41
chr1q	6610	0,04129263	5,04
chr10p	2174	0,01358096	-2,61
chr10q	4512	0,02818644	-4,14
chr11p	3047	0,01903459	-1,76
chr11q	4369	0,02729312	0,56
chr12p	2384	0,01489283	1,51
chr12q	6368	0,03978086	6,25
chr13p	0	0	0
chr13q	5922	0,0369947	-0,96
chr14p	0	0	0
chr14q	4494	0,02807399	-1,50
chr15p	0	0	0
chr15q	3276	0,02046515	-2,45
chr16p	1453	0,00907688	-2,01
chr16q	1887	0,01178808	-1,16
chr17p	855	0,00534118	-0,86
chr17q	1868	0,01166938	-0,24
chr18p	900	0,00562229	4,05
chr18q	3356	0,02096491	-1,28
chr19p	541	0,00337962	-0,16
chr19q	1434	0,00895819	4,24
chr2p	4505	0,02814271	-2,03
chr2q	9003	0,05624168	-1,25
chr20p	1312	0,00819606	-1,43
chr20q	1147	0,0071653	4,29
chr21p	215	0,0013431	-0,46
chr21q	2192	0,01369341	-2,11
chr22p	0	0	0
chr22q	819	0,00511629	-0,49
chr3p	4586	0,02864871	-3,63
chr3q	7069	0,04416	-1,69
chr4p	3076	0,01921575	-1,36
chr4q	9887	0,06176403	-2,27
chr5p	3247	0,02028399	-1,75
chr5q	8639	0,05396778	0,36
chr6p	2911	0,018185	-2,73
chr6q	7409	0,04628398	-2,72
chr7p	3180	0,01986544	2,55
chr7q	6116	0,03820661	2,15
chr8p	2248	0,01404324	-2,10
chr8q	6510	0,04066793	-2,20
chr9p	2669	0,01667323	-0,69
chr9q	3326	0,0207775	-0,05
chrXp	2720	0,01699182	9,75
chrXq	6394	0,03994328	11,26
chrYp	218	0,00136184	-3,87
chrYq	246	0,00153676	-2,61