Molecular screening for the presence of large deletions or duplications in *BRCA* using Multiplex Ligation-dependent Probe Amplification in South Africa

By

Pakiso James Moeti

2008046928

Submitted in fulfilment of the requirements in respect of the Magister Scientiae in Medical Sciences degree qualification (M.Med.Sc) in the Faculty of Health Sciences, Division of Human Genetics, University of the Free State, Bloemfontein, South Africa

Supervisor: Dr NC van der Merwe

Co-Supervisor: Me BK Dajee

July 2016

Declaration

I declare that the master's research dissertation or interrelated, publishable manuscripts / published articles that I herewith submit at the University of the Free State, is my independent work and that I have not previously submitted it for a qualification at another institution of higher education.

I hereby declare that I am aware that the copyright is vested in the University of the Free State.

I hereby declare that all royalties as regards to intellectual property that was developed during the course of and/or in connection with the study at the University of the Free State, will accrue to the University.

_____

Pakiso Moeti

Dedication to my anchors

First and foremost, this dissertation is dedicated to my parents for bringing me into the universe and grandparents for being the pillars of my strength, I vehemently thank you for contributing and making me the man I am today. My partner Tshidi, I thank you for believing in me and for your dedicated support and for your unconditional love.

Acknowledgements

It is indeed true that the writing and compilation of this dissertation was guided by key individuals, and without them it would not have been possible.

In particular, my deepest gratitude is to my study-leaders Dr NC van der Merwe and Ms B Dajee. Your enormous contributions regarding guidance, support, patience and motivation meant the world to me.

Sincere appreciation goes to the breast cancer patients for participation in the study. Without you, none of this would have been possible.

I wish to acknowledge and express my deepest gratitude towards the following institutions for financial and technical support:

- the Division of Human Genetics at the University of the Free State
- the National Health Laboratory Services (NHLS) for the use of facilities and resources
- the Medical Research Council of South Africa (MRC) for financial assistance
- the University of the Free State postgraduate school for financial assistance.

Family and friends for support and love.

Finally, the Creator and Almighty God who carried and blessed me throughout.

Summary

Germline *BRCA* gene mutations are associated with hereditary breast and ovarian cancers (OVC). Identification of these mutations greatly improves the preventive strategies and management of patients affected with the disease. The large majority of alterations identified within *BRCA1* and *BRCA2* are point mutations and small insertions/deletions. However, an increasing number of large genomic rearrangements (LGRs) are internationally being reported. Their contribution to familial breast cancer risk varies for different populations, for in some countries it represents a founder mutation (such as the Netherlands), whereas in others this type of mutation is totally absent. The main objective was to optimize and validate this new technique for use within the diagnostic laboratory and to screen various South African (SA) population groups for the presence of these larger genomic rearrangements present within *BRCA1* and *BRCA2*.

A total of 129 patients, who tested negative for the presence of smaller pathogenic *BRCA1* or *BRCA2* mutations were included in the study. The patients represented the Black, Indian and Coloured populations of South Africa. The selection criteria included being affected with breast cancer, have a minimum of one other family member affected with the disease or an early age at onset (diagnosed before the age of 45). Genomic DNA was extracted from peripheral blood samples. Multiplex Ligation-dependent probe amplification (MLPA) was performed using the SALSA® MLPA® probemixes P002-C1, SALSA® MLPA® P002-D1 and SALSA® MLPA® P087-C1 for *BRCA1* and SALSA® MLPA® probemixes P045-B3 and SALSA® MLPA® P077-A3 for *BRCA2*. The data obtained were analyzed by using the GeneMarker® software v 2.6.4.

Screening for the presence of LGRs within *BRCA1* and *BRCA2*, did not reveal any genomic rearrangements present within these genes. Although no patients were identified that carried this type of deletions or duplications, the use of the five different probe sets (two screening probe sets and one confirmation set for *BRCA1*

and two for *BRCA2*, of which one represented the confirmation set) were successfully validated for use on the diagnostic platform.

The data furthermore highlighted the dramatic effect that small deletions or duplications within these genes might have when situated within the critical ligation site of the specific probe set. The presence of these smaller mutations could result in false positive results.

The results of this study serve as a warning to pathology laboratories within SA, as the most common Afrikaner founder mutation situated within *BRCA2* exon 17 affects the ligation of the probe set for exon 17. The presence of this mutation resulted in a reduced signal for exon 17, therefore a false positive result. This places emphasis on the confirmation of all potential positive results by using an alternative method or different probemix in order to prevent reporting of a false positive result.

The data gathered corresponded to that of previous SA studies and supported the tentative hypothesis that LGRs do not seem to play a significant role within the various SA populations. It does not contribute significantly to the familial BC risk within SA.

Keywords: South Africa, familial breast cancer, large genomic rearrangements, Multiple ligation dependent probe amplification, mutations*, BRCA1, BRCA2.*

Opsomming

Oorerflike mutasies in die *BRCA* gene word geassosieer met familiële bors en ovariële karsinoom. Die identifisering van hierdie tipe mutasies kan voordele vir die aangetaste pasiënt inhou, rakende voorkomende strategieë en behandeling. Alhoewel die meerderheid van hierdie veranderinge in *BRCA1* en *BRCA2* enkel basispaar mutasies en klein invoegings of delesies is, word al hoe groter herrangskikkings in die genoom al meer in die internasionale literatuur beskryf. Die bydrae wat hierdie groter herrangskikkings maak tot die algehele oorerflike borskanker risiko, variëer. In sekere lande (soos Nederland) verteenwoordig die groter herrangskikkings 'n stigterseffek, terwyl dit feitlik afwesig is in ander. Die doel van die studie was om die gebruik van die nuwe tegniek (Multiplex Ligation-dependent probe amplification of MLPA) te optimiseer en te valideer, sodat dit met vertroue gebruik kan word om pasiënte van die Suid-Afrikaanse (SA) populasies te sif vir die teenwoordigheid van hierdie tipe mutasies binne *BRCA1* en *BRCA2*.

'n Totaal van 129 borskanker pasiënte is ingesluit in hierdie studie. Hierdie pasiënte het negatief getoets vir die teenwoordigheid van kleiner siekte-veroorsakende veranderinge in hierdie twee gene. Die pasiënte was verteenwoordigend van die Swart, Indiër en Kleurling bevolking van SA. Die pasiënt moes aangetas wees met borskanker, ten minste een ander aangetaste familielid in die familie hê of gediagnoseer wees voor ouderdom 45. Genomiese DNA is geëkstraheer vanuit volbloed. Die MLPA tegniek is uitgevoer deur gebruik te maak van vyf verskillende stelle peilstukke, drie vir *BRCA1* en twee vir *BRCA2* (SALSA® MLPA® P002-C1, SALSA® MLPA® P002-D1 en SALSA® MLPA® P087-C1 vir *BRCA1*; en SALSA® MLPA® P045-B3 en SALSA® MLPA® P077-A3 vir *BRCA2*). Die data is verwerk deur gebruik te maak van GeneMarker® sagteware (weergawe 2.6.4).

Sifting vir die teenwoordigheid van LGRs binne *BRCA1* en *BRCA2*, het geen positiewe resultate opgelewer nie. Hoewel geen pasiënte geïdentifiseer is wat oor groter herrangskikkings beskik nie, is die gebruik van die vyf verskillende

ondersoekstelle suksesvol ge-optimiseer en ge-implementeer vir gebruik in die diagnostiese laboratorium.

Die studie beklemtoon egter die drastiese effek wat klein delesies of duplikasies binne hierdie gene kan hê wanneer die spesifieke mutasie in die ligeringsgebied van die peilstukke geleë is. Die teenwoordigheid van hierdie kleiner mutasies kan tot vals positiewe MLPA resultate lei.

Die resultate van hierdie studie dien as 'n waarskuwing aan patologie laboratoriums binne SA, aangesien die mees algemene Afrikaner stigtersmutasie in *BRCA2* ekson 17 in so ‘n gebied geleë is. Die teenwoordigheid van hierdie mutasie in die ligeringsgebied lei tot ’n verlaging in die sein vir ekson 17, met ander woorde ’n vals positiewe MLPA uitslag. Dit plaas klem op die feit dat alle positiewe MLPA resultate bevestig moet word deur gebruik te maak van ’n alternatiewe metode of ’n tweede ondersoek stel. Sodoende sal foutiewe positiewe uitslae voorkom word.

Die data verkry uit hierdie studie stem ooreen met die van vorige SA studies en ondersteun die tentatiewe hipotese dat LGRs nie ‘n belangrike rol in die verskillende Suid-Afrikaanse bevolkings speel nie. Die teenwoordigheid van hierdie tipe herrangskikkings dra nie beduidend by tot die familiële risikos vir oorerflike borskanker in SA nie.

Sleutelwoorde: Suid-Afrika, oorerflike borskanker, groter genomiese herrangskikkings, Multiple ligation dependent probe amplification (MLPA), mutasies, *BRCA1, BRCA2.*

Contents

List of Figures

List of Tables

Abbreviations and acronyms

®	Registered Trademark
™	Trademark
5’	Five prime end
3’	Three prime end
α	Alpha
β	Beta
°C	Degree Celsius
Ψ	Pseudogene
µl	Microliter
A	Adenine
aa	Amino Acids
Ala	Alanine
Arg	Arginine
Asp	Aspartic acid
ATM	Ataxia Telangiectasia
bp	Base Pair
BACH1	BRCA1-Associated Carboxyl-Terminal Helicase
BAP1	BRCA1 Associated Protein
BARD1	BRCA1-Associated RING Domain
BART	BRCAnalysis Rearrangement Test
BC	Breast Cancer
BIC	Breast Cancer Information Core

BRC	Breast Cancer Repeats Motifs
BRCA1	Breast Cancer susceptibility gene 1
BRCA2	Breast Cancer susceptibility gene 2
BRIP1	BACH1-BRCA1-Associated C-terminal Helicase
BRCT	BRCA1 Carboxyl Terminus
c	Coding
C	Cytosine
ca	Cancer
cDNA	Complimentary DNA
CHEK2	Checkpoint Kinase 2
COA	Certificate of Analysis
C-terminus	Carboxyl Terminus
Cys	Cysteine
DBD	DNA Binding Domain
DEAH	Helicase box
Del	Deletion
dH_2O	Distilled Water
DNA	Deoxyribose Nucleic Acid
dNTP	Deoxynucleotide Triphosphate
DQ	Dosage Quotient
dsDNA	Double-stranded DNA
DSS1	Deleted in Split-hand/Split-foot 1 region
DTT	Dithiothreitol
Dup	Duplication
dx	Age of onset
ECUFS	Ethics Committee of the University of the Free State

EDTA	Ethylenediaminetetraacetic Acid
ER	Estrogen Receptor
FANCD1	Fanconi Anaemia Gene 1
FISH	Fluorescent *in situ* Hybridization
g	Gravitational force
G	Guanine
Gln	Glutamine
Glu	Glutamate
Gly	Glycine
GS	Gene Scan
HCl	Hydrochloric Acid
HDAC	Histone Deacetylase
Hi-Di	Highly Deionized Formamide
His	Histidine
HRMA	High Resolution Melting Analysis
IHC	Immuno-histochemical
Ile	Isoleucine
Ins	Insertion
KCl	Potassium Chloride
kb	Kilobase
kDa	Kilo Dalton
kV	Kilovolt
LFS	Li-Fraumeni Syndrome
LGRs	Large genomic Rearrangements

LOH	Loss of Heterozygosity
Lys	Lysine
$MgCl_2$	Magnesium Chloride
MIM	Mendelian Inheritance of Man
ml	Millilitre
MLPA	Multiplex Ligation-dependent Probe Amplification
mM	Millimolar
mRNA	Messenger RNA
NaCl	Sodium Chloride
NCBI	National Centre for Biotechnology Information
$ng.\mu l^{-1}$	Nanogram per Microliter
NGRL	National Genetics Reference Laboratories
NGS	Next Generation Sequencing
NHLS	National Health Laboratory Services
NLS	Nuclear Localization Signal
nm	Nanometre
nt	Nucleotide
N-terminus	Amino Terminus
NBR1	Next BRCA1
OB	Oligonucleotide/oligosaccharide-Binding
OVC	Ovarian Cancer
p	Protein
PALB2	Partner and Localizer of BRCA2
PCR	Polymerase Chain Reaction
pH	Potential of Hydrogen

Phe	Phenylalanine
PR	Progesterone Receptor
Pro	Proline
PTEN	Phosphatase and Tensin Homolog
RAD51	Homology of RecA of E.coli
RHA	RNA Helicase A
RNA	Ribonucleic Acid
RPA	Relative Peak Area
rs	Reference SNP
SA	South Africa
SDS	Sodium Dodecyl Sulphate
Ser	Serine
SET	Sodium Chloride EDTA Tris-HCl
SINE	Short Interspersed Nuclear Elements
SSCP	Single Stranded Confirmation Polymorphism
ssDNA	Single-stranded DNA
STK11	Serine/Threonine Kinase 11
T	Thymine
Thr	Threonine
Tris	2-amino-2-(hyroxymethyl)-1,3-propanediol
Tyr	Tyrosine
UFS	University of the Free State
US	United State
UTR	Untranslated Region
v/v	Volume per Volume

Val	Valine
WHO	World Health Organization
w/v	Weight per Volume

Chapter 1

Literature Review

1.1 Introduction

South Africa (SA) is experiencing a dramatic rise in infectious and non-communicable diseases such as cardiovascular diseases, type 2 diabetes and now also cancer (Mayosi *et al*., 2009). It is reaching epidemic proportions, especially in typical low-income countries such as SA. Urbanization of the previously disadvantaged ethnic groups gave rise to a higher social status especially for the Black and Coloured populations. This brought about a change in diet, lifestyle and other environmental factors that made them more susceptible to various diseases. This has resulted in an increase to the number of cancer cases reported (Somdyala *et al*., 2010).

Since the discovery of the two hereditary breast cancer (HBC) genes, considerable progress has been made in the characterization of the genetic component of BC (Tonin *et al.*, 1996; Sokolenko *et al.*, 2006). For SA, the investigations have been complicated due to the unique genetic backgrounds or heritage for each of the ethnic groups. Various research projects have been aimed at elucidating the role of *BRCA1* (Breast cancer susceptibility gene one– Online Mendelian Inheritance of Man [OMIM] 113705) and *BRCA2* (Breast cancer gene number two – OMIM 600185) within the Caucasian Afrikaner population (Reeves *et al.,* 2004; Schlebusch *et al*., 2010), as females representing this population group has the highest risk of developing the disease (one in 13 as indicated by Vorobiof *et al.,* 2001 and Reeves *et al.,* 2004).

Although these results, together with that of various pilot studies performed for the Black, Indian and Coloured populations were met with high expectations and resulted in the establishment of a pathology driven diagnostic service for these

groups, the research was not complete. A single outstanding research question still remained, namely to what extent does the presence of larger genomic rearrangements (frequently reported in the international literature) contribute to the familial BC burden in SA population groups other than the Afrikaner? This question formed the basis of the research presented here and involved the SA Black, Indian and Coloured populations (Coloured is commonly used which refer to a South African community of mixed ancestry)(Chimusa *et al*., 2011).

1.2 Breast Cancer

The advancement of diagnostic techniques and treatment in the last decade has greatly contributed to the survival of cancer patients. Although this is true, BC remains one of the leading causes of death in women today (Ferlay *et al.*, 2015). It is characterized as a malignant tumor of breast tissue suspected by clinical findings such as a breast lump, breast thickening, skin changes or density changes visible using mammography. The cancer is staged from 0 (earliest) to IV (most advanced) with survival being dependent upon the stage of diagnosis.

Breast cancer is the most frequently diagnosed cancer and the leading cause of cancer death among females worldwide, with an estimated 1.7 million cases and 521,900 deaths in 2012 (Torre *et al.*, 2015). BC alone accounted for 25% of all international reported cancer cases during that year and 15% of all cancer deaths among females. Although the age-standardised rate per 100 000 was relatively low for Southern Africa compared to the developed countries situated within Western Europe, Northern America and Northern Europe, the mortality rate was similar for these regions (Fig. 1.1).

In less developed regions of Africa, BC is the second leading cause of death accounting for 11.0% after lung cancer with 13.3% (Ferlay *et al*., 2015). The highest rates are seen in countries such as Egypt, Nigeria, Algeria, and SA (Parkin *et al.*, 2014). According to the World Health Organization (WHO), cancer is increasing in Africa. It was hypothesized that this is due to the aging and growth of the population, limited resources, as well as increased prevalence of risk factors mainly associated with economic transition (Parkin *et al.*, 2014). The increase is specifically linked to urbanization and economic developments within these countries (Parkin *et al*., 2014).

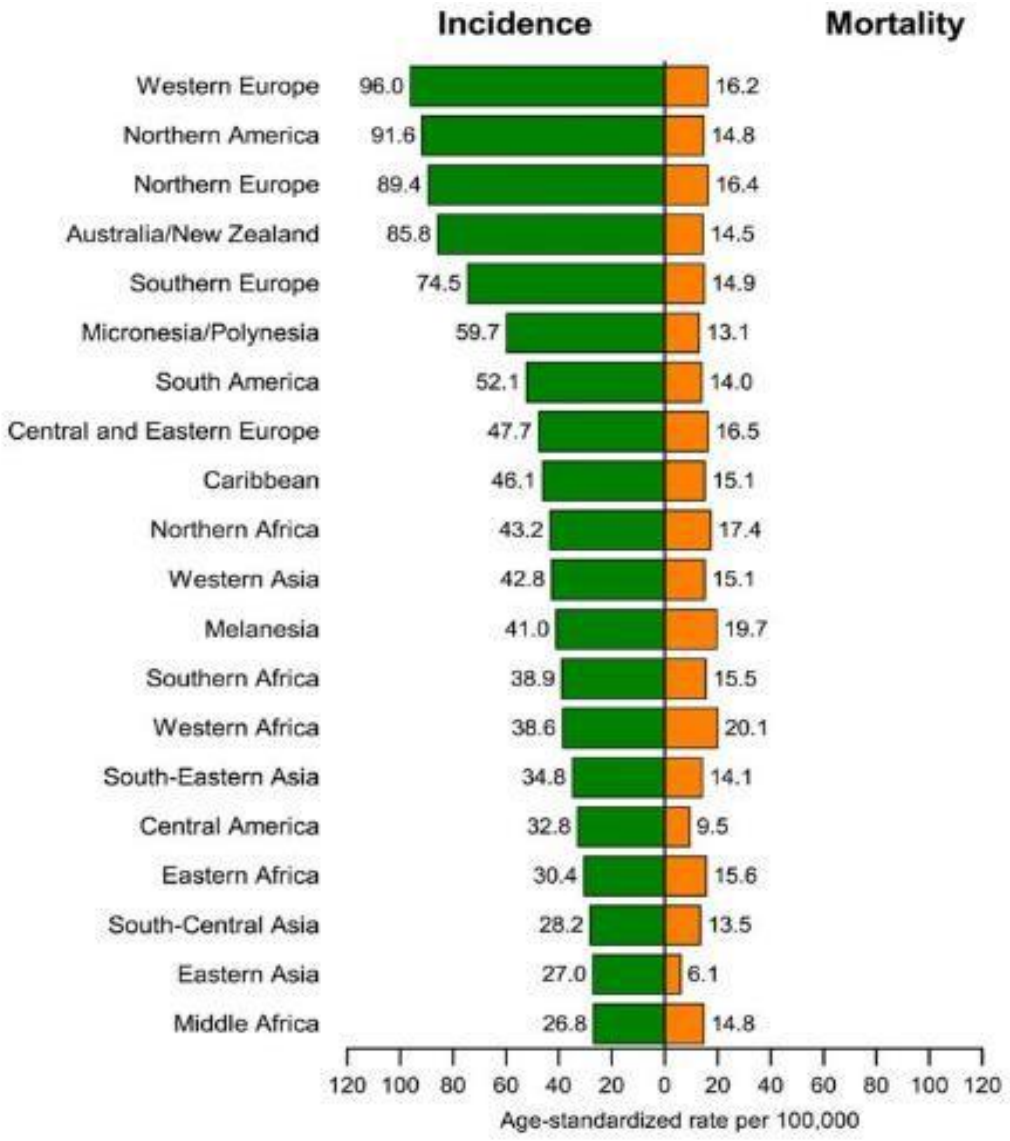

Figure 1.1 BC incidence and mortality rates according to different world areas (Torre *et al.*, 2015). Included in Southern Africa are SA, Namibia and Zimbabwe (accessed online on December 1, 2015).

Despite the increase in incidence, cancer continues to receive a relatively low public health priority in Africa despite the growing burden. The disease remains undiagnosed until it is late or in a metastatic stage when the treatment options have less benefit or are simply unavailable (Anderson *et al*., 2011).

In SA, the source of information on cancer incidence and mortality is the National Cancer Registry, which collects information from pathology laboratories on histologically diagnosed cancers (Vorobiof *et al.*, 2001, National Cancer Registry, 2010). Based on their statistics, the same scenario is also applicable to SA, with BC being the most common malignancy affecting women. The country has a crude incidence rate of 18.5/100 000 based on cases recorded between 1993 and 1995 (Schlebusch *et al.*, 2010). Of these patients, a small but significant percentage (5 - 10%) is considered to be directly due to an inherited susceptibility (Diamond *et al.*, 1998; Liu & West, 2002; Thompson & Easton, 2002; Sigurdson *et al*., 2004).

Approximately one in every 22 SA women is at risk to develop the disease (SA National Cancer Registry, 2010). The lifetime risk varies for the different SA population groups, with a lifetime risk of one in 49 for Black woman, one in 18 for Coloured women to one in 13 for the Caucasian female population, which includes the Afrikaner (Vorobiof *et al*., 2001; Reeves *et al*., 2004). The variation in the risk might be attributed to various epidemiological factors such as reproductive factors (nulliparity, early age at menarche, late age at menopause, late age at first full-term pregnancy and breastfeeding), physical inactivity and the presence of founder effects within specific populations. The disease is ranked the most common cancer diagnosed in Caucasian (17.9%) and Asian (24.4%) patients and is the second most common cancer in the Coloured (18.2%) and Black (13.4%) women (Sitas *et al*., 1998). BC in SA is associated with a high mortality rate, mainly as a result of delayed diagnosis, for reasons that include limited community awareness and restricted access to oncology care facilities at provincial hospitals (Schoeman *et al.*, 2013).

1.3 Familial breast cancer

Hereditary (familial) BC occurs when a family is suggestive of an inherited predisposition. The inheritance is typically a dominant trait where more than one member within a family is affected. This type of inheritance characteristically accounts for a small but significant estimated proportion of all BC cases (Easton *et al*., 1995; Claus *et al*., 1996).

The congenital susceptibility for the development of BC is partly linked to the inheritance of mutations in the familial BC genes *BRCA1* and *BRCA2*. These two tumor suppressor genes have been mapped and cloned a couple of years apart [*BRCA1* by Miki *et al*. (1994); *BRCA2* by Tavtigian *et al.* (1996)]. Together, they explain 20 - 40% of heritable BC cases in various populations over the world (Wooster & Weber, 2003; Thompson & Easton, 2004). Statistics show that a woman who carries a mutated copy of *BRCA1*, has an up to 85% risk of developing BC by age 70 (Tutt & Ashworth, 2002; Wooster & Weber, 2003), compared to a 12% risk in the general population (Burke & Austin, 2002).

The *BRCA* genes are tumor suppressor genes encoding proteins whose normal function is to inhibit cell transformation. Their inactivation typically occurs when the genes are mutated and this is advantageous for tumor growth and survival. This is in accordance with Knudson's (1971) 'double hit hypotheses' where the wild-type allele of the gene is lost in tumors of heterozygous carriers. Individuals carrying a defective gene copy are consequently predisposed to carcinomas of the breast and various associated cancer types.

Apart from the high risk *BRCA* genes, several other genes are also associated with an inherited susceptibility to BC. These additional genes are usually characterized by the penetrance of the disease. These include the BC syndromes such as Li-Fraumeni [LFS - OMIM 151623 caused by mutations within tumor suppressor p53 gene *Tp53* (OMIM 191170)] and Cowden syndrome [OMIM 158350 caused by Phosphatase and tensin homolog gene *PTEN* (OMIM 601728)] (Li & Fraumeni 1969;

Liaw *et al*., 1997). These two genes, together with several others are considered to have a moderate penetrance to BC. Included into this group is the Ataxia Telangiectasia gene (*ATM* - MIM 208900), serine/threonine kinase gene 11 (*STK11* - OMIM 175200) which causes Peutz-Jeghers syndrome when mutated (Lehur *et al*., 1984; Mehenni *et al*., 1998). The low to moderate penetrance variants include a Checkpoint Kinase 2 (*CHEK2* - OMIM 604373) (Meijers-Heijboer *et al*., 2002), Partner and Localizer of BRCA2 (*PALB2* - OMIM 610355) and BRCA1 Interacting Protein C- terminal Helicase 1 (*BRIP1* - OMIM 605882).

1.3.1 *BRCA1*

BRCA1 is a large gene positioned on chromosome 17 comprising of 24 exons, of which 22 are coding (Fackenthal & Olopade, 2007). The largest exon is exon 11 coding for approximately 60% of the protein (Miki *et al*., 1994 and Chen *et al*., 1995). The entire coding region consists of 5589 nucleotides, which is transcribed into an mRNA of 7.8 kilobase (kb). The protein consists of 1863 amino acids (aa), with a molecular weight of 220 kilodalton (kDa) (Chen *et al*., 1995).

BRCA1 shows no sequence homology to any other gene. Five transcript variants have been described (http://www.ncbi.nlm.nih.gov/gene/672) with *BRCA1* transcript variant 1 (NM_007294.3, coding sequence: 233 - 5824 base pairs [bp]) representing the most abundant transcript which encodes the full-length protein. This sequence has been incorporated as the reference standard in the NCBI RefSeqGene project. The ATG translation start site is located in exon 2 (known as the initiation site) (233 - 235 bp), with the stop codon situated in exon 24 (termination site) (bp 5822 - 5824). This reference sequence is derived from the U14680.1 sequence comprising of 5711 nucleotides, with a coding sequence ranging from bp 120 - 5711. This version includes longer 5' and 3' untranslated regions. The other four variants are rare variants that use alternative transcription sites (exon 1b) and/or alternative in-frame splice sites in the coding sequence. These smaller alternatively spliced transcripts all have distinctively different patterns of expression (Orban & Olah, 2003).

Transcription starts from one of two alternatively spliced promoter regions (*BRCA1* 1a or α- and 1b or β-) to yield two distinct transcripts (Xu *et al*., 1995). The two promoters control the expression of these two transcripts, which regulate transcription, translation and the alternative splicing of *BRCA1* (Xu *et al.*, 1997). According to Xu and co-workers (1995), *BRCA1* exon 1a is more efficiently translated, as this exon has no upstream AUG codon in the 5'UTR of the mRNA in contrast to exon 1b, which has an upstream AUG codon (Xu *et al*., 1995). Each of the alternative first exons is linked by splicing of exon 2 at an RNA level (Pamula *et al.*, 2006). Translation, however starts in exon 2.

A BRCA1 pseudogene, which is an imperfect version of the functional gene, has high sequence similarity to *BRCA1* exons 1a, 1b, and 2, located 40 kb upstream of exon 1 (Brown *et al.*, 1996). This pseudogene is formed by the partially duplicated 5' ends of both *BRCA1* and the gene that lies downstream of *BRCA1*, namely *NBR1* (Next to *BRCA1,* also known as 1A1-3B) (Fig. 1.2). This partial copy of the *BRCA1* gene (Ψ-*BRCA1*) consists of the duplicated exons 1a, 1b and 2 of *BRCA1* and lies next to a partial copy of the *NBR1* gene (termed *NBR2*) (Fig. 1.2). The partial duplication of *NBR1* consists of three exons (exons 1a, 1b and 3), together with 295 bp of the intergenic region (Brown *et al*., 1996). The intron-exon structure is maintained in both the pseudogenes Ψ-*BRCA1* and *NBR2*, which shows a high degree of nucleotide sequence identity between the respective genes. This suggests that these genes are non-processed pseudogenes (Brown *et al*., 1996; Xu *et al.*, 1997).

The BRCA1 protein is normally located in the nucleus and contains phosphorylated residues (Chen *et al*., 1996). Two recognizable protein motifs are found, namely the RING finger domain and the two BRCA C-terminus (BRCT) repeat domains. The highly conserved RING finger motif is situated near the amino terminus (N-terminus) (Miki *et al.*, 1994) (Fig. 1.3). This motif consists of a zinc binding domain that includes a conserved pattern of seven cysteines and one histidine (Miki *et al.*, 1994). The BRCA1 RING finger facilitates protein-protein and protein-DNA interactions by specifically binding with both BAP1 (BRCA1-associated protein) and BARD1 (BRCA1-associated RING domain), another RING finger protein (Miki *et al*., 1994; Wu *et al.*, 1996; Brzovic *et al*., 2001) (Fig. 1.3). Together, these two genes act in

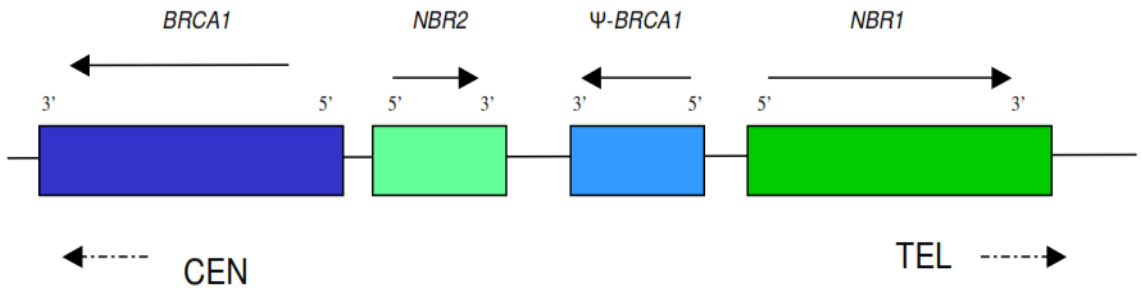

Figure 1.2 Illustration of the *BRCA1-NBR1* region. Solid arrows indicate the direction of transcription, whereas dashed arrows indicate the position of these genes with regards to the centromere and the telomere. Adapted by Reeves (2006) from Xu *et al.* (1997).

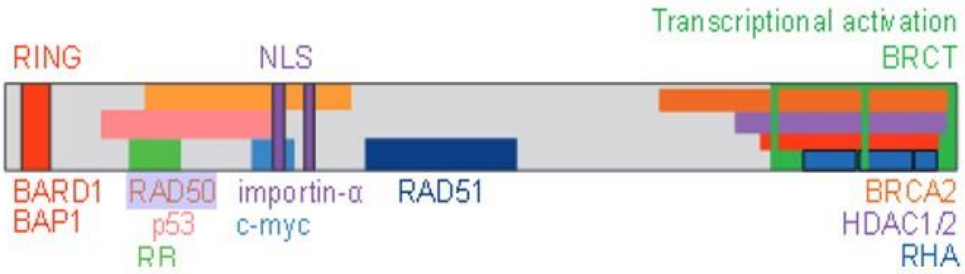

Figure 1.3 Structural and functional features of the BRCA1 protein (total of 1863 aa). The N-terminal RING finger interacts with BRCA1-associated RING domain 1 protein (BARD1) and with the deubiquitinating enzyme, BRCA1-associated protein-1 (BAP1). Two nuclear localization signals (NLSs) bind importin-a; nearby regions interact with p53 (TP53), retinoblastoma protein (RB), RAD50, and MYC. A domain within BRCA1 aa 758–1064 interacts with RAD51. The BRCA1 C-terminal (BRCT) repeats interact with BRCA2, histone deacetylase (HDAC) 1 and 2, RNA helicase A (RHA), and the CtBp-interacting protein (CtIP). The putative transcriptional activation domain lies at the C-terminus (Welcsh *et al.*, 2000).

promoting tumor suppression (Simon *et al.*, 2006).

The protein also has two nuclear localization signals (NLS), which are situated in exon 11 (Fig. 1.3). These two NLS represents a central transcription activation domain which is able to bind DNA, and to an RAD51 binding domain (Fig. 1.3). Both of these are suggestive of a role in DNA repair.

Another functional domain is situated within the C-terminal of the gene (Fig. 1.3). This region is called the BRCT (*BRCA1* carboxyl-terminal) repeats and facilitates protein-protein interactions (Hohenstein & Fodde 2003). These repeats are each about 110 aa long and comprise aa 1653 - 1736 and 1760 - 1855. The BRCT domains are usually involved in DNA repair, transcriptional co-activation and cell cycle regulation. Williams and co-workers (2003) reported that more than 60% of clinically relevant *BRCA1* mutations delete a portion of or all of the BRCT domains, which highlights the critical role of the C-terminus. The majority of missense mutations in the BRCT that was found to influence the folding determinants of the domain resulted in the destabilization of the protein (Williams *et al.*, 2003). The BRCT domain of *BRCA1* directly interacts with BACH1 (BRCA1 associated carboxyl-terminal helicase) and with BRIP1 (BACH1-BRCA1-associated C-terminus helicase-1), both members of the DEAH helicase family.

BRCA1 plays a pivotal role in a number of super complexes involved in DNA damage response activation and double strand DNA repair. Due to its involvement, the protein interacts directly or indirectly with numerous molecules, ranging from DNA damage repair proteins, oncogenes, to cell cycle regulators, transcriptional activators, and repressors (Deng & Brodie 2000). Loss-of-function mutations occurring in *BRCA1* result in pleiotropic phenotypes, which can include consequences ranging from increased apoptosis, defective DNA damage repair and abnormal centrosome duplication, to chromosome damage (Brodie & Deng 2001; Deng 2002; Venkitaraman 2002). Based on these effects, it was proposed and concluded that mutations in *BRCA1* result not in tumor formation itself, but rather in genetic instability, subjecting cells to a high risk of malignant transformation (Lengauer *et al*., 1998; Deng 2001).

1.3.2 *BRCA2*

The *BRCA2* gene is located on chromosome 13q and was identified in 1994 as a possible second BC predisposing gene (Wooster *et al*., 1994). Its role in familial BC was confirmed within two years (Wooster *et al*., 1995; Tavtigian *et al*., 1996). This gene is even larger than *BRCA1* with very little homology between the two. The gene has an 11385 bp transcript which codes for a nuclear protein of 3418 aa (384kD). The NM_000059.3 sequence acts as the reference standard as used in the NCBI RefSeqGene project.

The gene is composed of 27 exons, of which 26 are coding. Translation also starts in exon 2. The only similarity that *BRCA2* has with *BRCA1*, is the large exon 11 situated in the middle of the gene. This exon 11 comprises more than 60% of the coding sequence and contains eight BRC (Breast cancer repeat) motifs, which is related to its function (Fig. 1.4). Although the structure of the gene is not as well characterised as *BRCA1*, it has been proven that these BRCs are conserved across mammals and consist of 30 - 80 aa each (Wong *et al.*, 1997). Four of these motifs interact with RAD51 (Wong *et al.* 1997) (Fig. 1.4). The role of these repeats is to mediate protein-protein interaction and to participate in DNA repair and recombination.

The gene also has two nuclear localization signals (NLSs) at the C-terminus, both of which are responsible for the binding of the gene to RAD51 (Fig.1.4) (Roy *et al*., 2012). These NLSs are situated within the final 156 residues of the gene and are essential for its cellular localization (Spain *et al*., 1999). Mutations predicted to prematurely truncate the BRCA2 protein 5' to the NLSs would render it cytoplasmic and rule out any interaction with the RAD51 complex unless it is transported into the nucleus by alternative means (Welcsh *et al.*, 2000; Roy *et al*., 2012). The C-terminal region of BRCA2 (~1000 residues), which include the NLSs are the most conserved portion of the protein (Warren *et al.*, 2002).

Studies have also indicated that the gene contains a DNA/DSS1 binding domain (BRCA2DBD) situated from aa 2478 - 3185. This group of aa forms a helix-turn-helix motif, which consists of an OB1 (oligonucleotide/oligosaccharide-binding), an

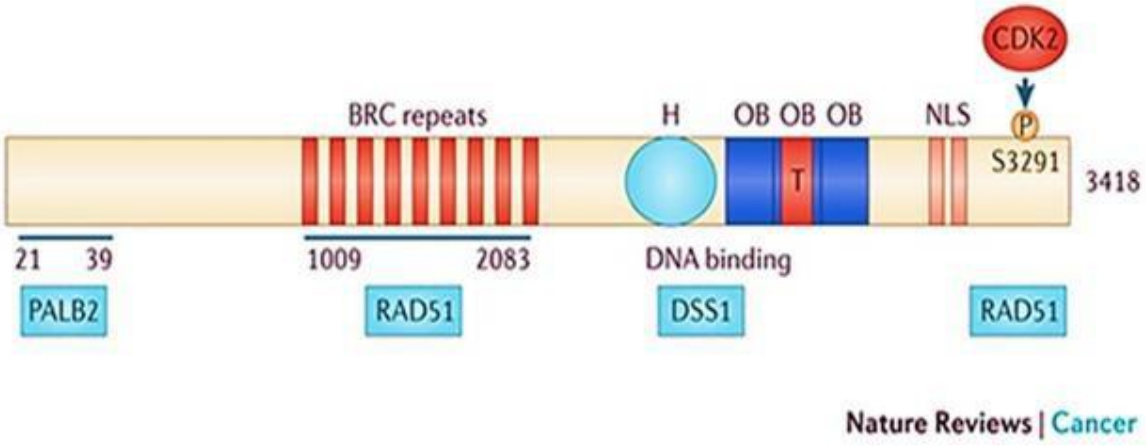

Figure 1.4 An illustration of BRCA2 functional domains and interaction with other proteins. Indicated are the eight BRC repeats and the two NLSs located at the C-terminus of the protein. The diagram also depicts various protein interactions (such as PALB2, DSS1 and RAD51) with each domain (Roy *et al*., 2012).

OB2, OB3 and a Tower region (Fig 1.4) (Yang *et al.*, 2002). Single-stranded DNA (ssDNA) binding has been attributed to two of the three OB folds, whereas the Tower region has been implicated in the interaction with double-stranded DNA (dsDNA). The presence of these domains implicates BRCA2 in both ssDNA and dsDNA binding (Yang *et al.*, 2002). *BRCA2* has also been identified as identical to one of the Fanconi anaemia genes (*FANCD1*) (Wagner *et al.*, 2004).

1.3.3 BRCA function

The two *BRCA* genes have initially been characterized as tumor suppressors with cancer inhibiting properties. Later research indicates that both *BRCA1* and *BRCA2* are involved in maintaining genome integrity at least in part by engaging in DNA repair, cell cycle checkpoint control and even the regulation of key mitotic or cell division steps (O'Donovan & Livingston, 2010). Due to their function in cell cycle regulation and the cellular damage response, mutations in these genes are expected to lead to susceptibility, resulting in functional deregulation and cancer in more than one tissue type. The genes require two mutations to lead to tumor development in cells according to Knudson's double hit hypothesis (Knudson, 1971). It is not completely understood why mutations in these two genes are mainly involved in malignancies of the breast and ovaries, but it is thought that the interaction with estrogen and progesterone could play a role (Schlebusch *et al.*, 2010).

1.4 Mutations within *BRCA1* and *BRCA2*

Mutations within the familial BC genes are without doubt important determinants of risk for breast and/or ovarian cancers, although they are not the only genes involved in familial BC (de Jong e*t al.,* 2002). Women with a three generation family history of breast and/or ovarian cancer that test negative for mutations within *BRCA1* and *BRCA2*, may have a mutation in an as yet undiscovered BC gene (Neuhausen, 2000). Various researchers believe that apart from *BRCA1* and *BRCA2,* there may be several other genes of possibly moderate to lower risk that could account for a proportion of non-*BRCA* BC (Lakhani *et al.*, 2000, Stratton & Rahman, 2008, Zhang *et al*., 2011b).

The phenotype of BCs in women carrying *BRCA1* mutations differs from that of women carrying *BRCA2* mutations and sporadic cases (Williams *et al.,* 2006). *BRCA1* mutations tend to be of higher histological grade, have a higher proportion of tubular differentiation, all of which are poor prognostic features (Williams *et al.,* 2006). *BRCA2* associated tumors are more similar to sporadic breast tumors. These tumors are more often of intermediate grade, are normally hormone receptor positive and occur at later ages compared to *BRCA1* associated tumors (Williams *et al.,* 2006).

Pathogenic germline mutations in these genes are characterized by an increased risk for BC, ovarian cancer (OVC), prostate cancer and pancreatic cancer (Lux *et al.*, 2006). The lifetime risk for these cancers in individuals with a pathogenic variant is 40 – 80% for BC, 11 – 40% for OVC, 1 – 10% for male BC and up to 39% for prostate cancer. The risk of developing pancreatic cancer ranges between 1 – 7% (Petrucelli *et al.*, 2013 http://www.ncbi.nlm.nih.gov/books/NBK1247/) accessed on December 17, 2015). Individuals who carry a *BRCA2* pathogenic variant are also at risk to develop melanoma.

1.4.1 Point mutations within *BRCA1* and *BRCA2*

All the various mutation types have been recorded for both genes. These include single base changes (missense, splice site, synonymous and nonsense mutations), single and small bp deletions and duplications (resulting in frameshift and possibly splice site mutations) and in frame insertions and deletions (Figs. 1.5 & 1.6). For both genes, pathogenic mutations are recorded throughout the entire coding region, therefore no mutational “hot spot” exists (Cipollini *et al*., 2004, Thompson & Easton, 2004) (Figs. 1.5 & 1.6). The highest mutation frequencies observed for *BRCA1* were for both frameshift (red lines) and missense mutations (green lines) (Fig. 1.5) compared to missense mutations (green lines) only for *BRCA2* (Fig. 1.6). Splice site, in frame deletions, and duplications were in the minority for both genes (Breast cancer information core (BIC)). For *BRCA2*, more synonymous mutations were detected compared to *BRCA1*. This might be due to the completion of the *BRCA1* Circos database, which includes functional data for the majority of *BRCA1* missense mutations.

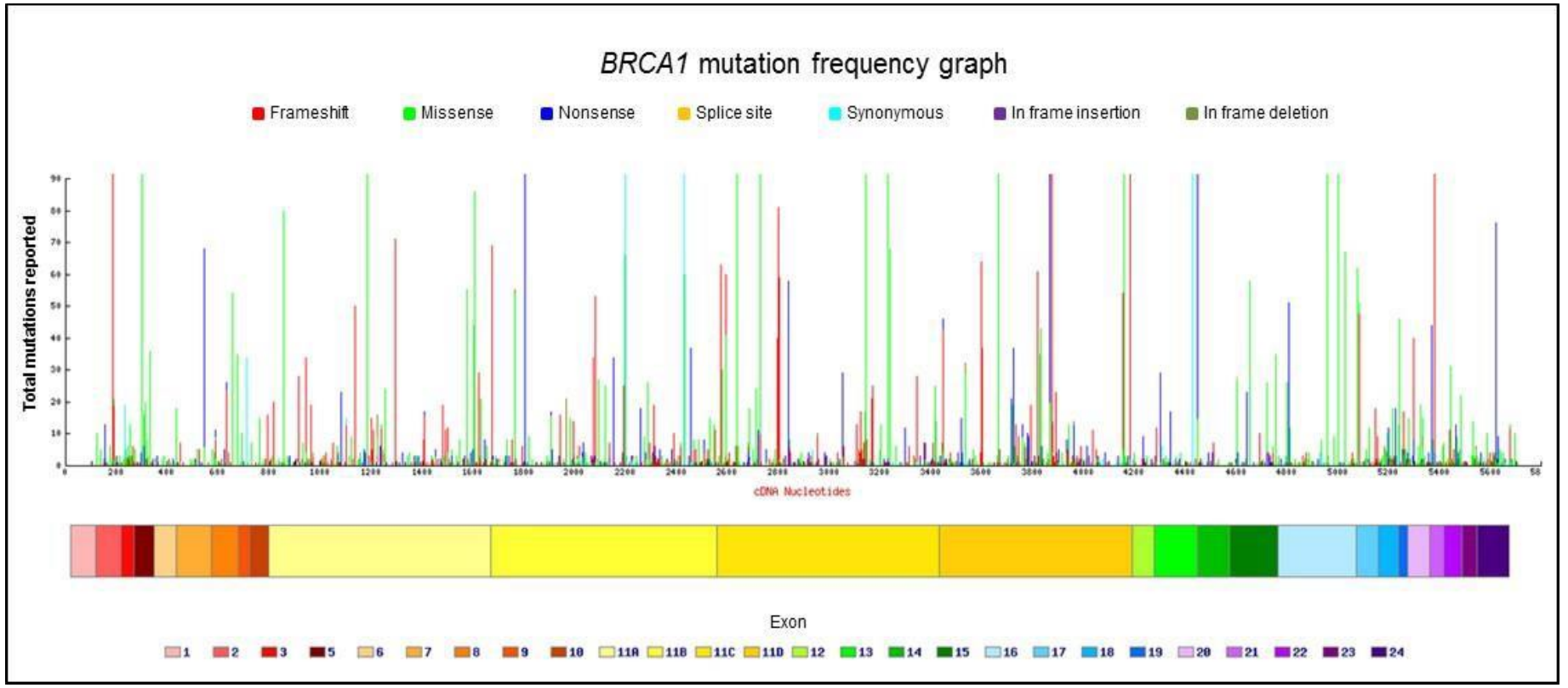

Figure 1.5 Frequencies of the types of mutations detected for *BRCA1* according to the Breast cancer Information Core (BIC). Indicated is the mutation type, the number of mutations reported for that specific exon, together with the cDNA nucleotide number (http://research.nhgri.nih.gov/bic/, accessed on December 28, 2015).

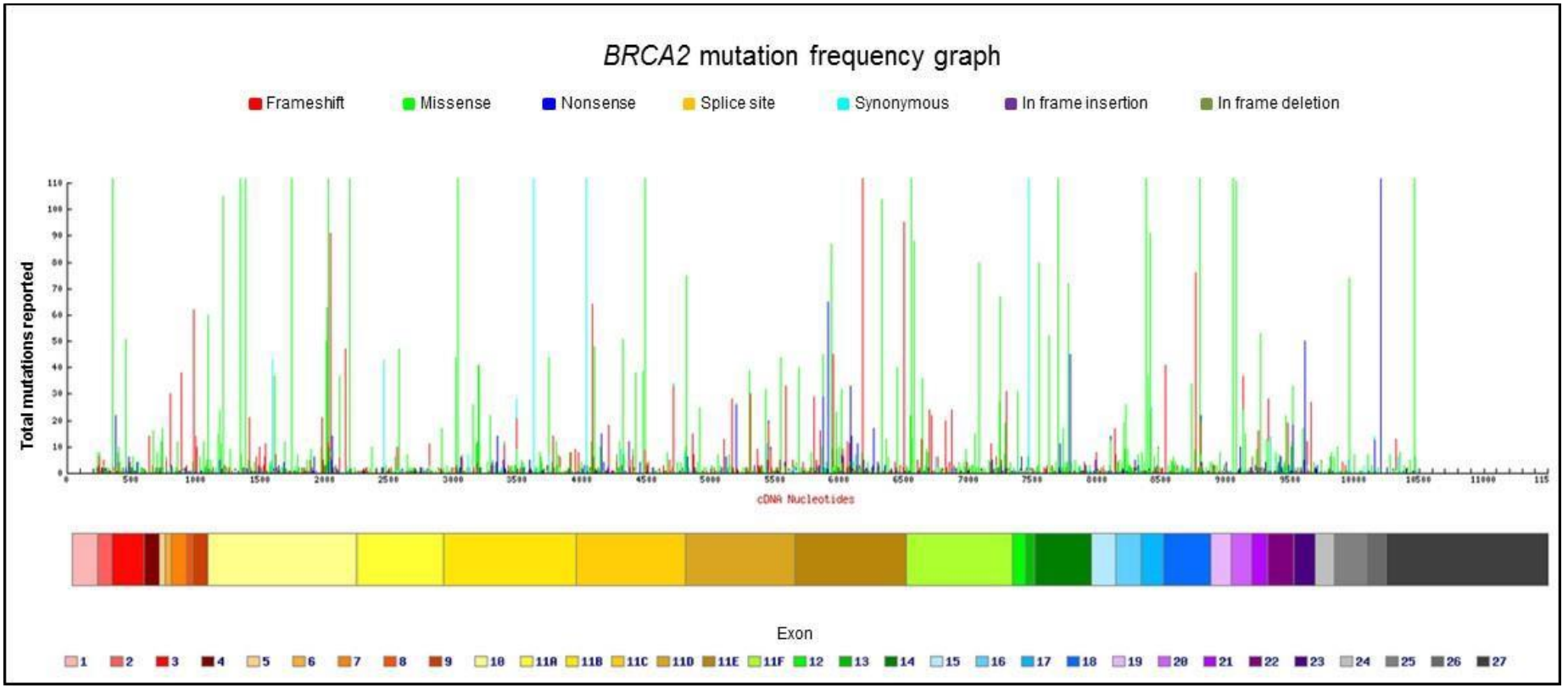

Figure 1.6 Frequencies of the types of mutations detected for *BRCA2* according to the BIC. Indicated is the mutation type, the number of mutations reported for that specific exon, together with the cDNA nucleotide number (http://research.nhgri.nih.gov/bic/, accessed on December 28, 2015).

1.4.2 *BRCA1* and *BRCA2* founder mutations in South Africa

Founder mutations are very common within the SA population, especially for the Afrikaner due to the presence of high linkage disequilibrium (Hall *et al.*, 2002). Founder mutations have been described for many genetic diseases ranging from Variegate porphyria (Dean, 1963) to progressive familial heart block type 1 (Torrington *et al.*, 1986). The presence of a founder effect within a population implies the loss of genetic variation that occurred with the establishment of a new population by a very small number of individuals representing a larger population (http://wallace.genetics.uga.edu/groups/evol3000/wiki/fb221/Bottlenecks_and_Founder_Effects.html, accessed on November 13, 2015).

The best-known example of a founder effect is present within the Ashkenazi Jewish population. This population group has been studied extensively throughout the years and is very well described where familial BC is concerned. This group of people are descendants of ancestors from Eastern and Central Europe. The most well characterized three founder mutations are *BRCA1* c.68_69delAG, p.Glu23ValfsX17 (BIC: 185delAG), *BRCA1* c.5266dupC,p.Gln1756ProfsX74 (BIC: 5382insC) and *BRCA2* c.5946delT,p.Ser1982ArgfsX22 (BIC: 6174delT) (Friedman *et al.*, 1995; Neuhausen *et al.*, 1996; Tonin *et al.*, 1996). These three founder mutations account for 98-99% of identified mutations within this population group and are carried by approximately 2.6% (1/40) of the Ashkenazi Jewish population (1%, 0.13% and 1.52% respectively) (Roa *et al.*, 1996; Frank *et al.*, 2002; Phelan *et al.*, 2002).

For SA, only a few founder mutations have been described thus far, with a total of three for the Afrikaner and a single mutation detected for the Coloured and Xhosa populations from the Western Cape (van der Merwe *et al.*, 2012). Research performed for the Afrikaner revealed the presence of three *BRCA* founder mutations [*BRCA1* c.1374del,p.Asp458GlufsX17 (1493delC), *BRCA1* c.2641G>T,p.Glu881X (2760G>T, E881X) and *BRCA2* c.7934del,p.Arg2645AsnfsX3 (8162delG)] (van der Merwe *et al.*, 2012). Together, these mutations account for the majority (≥ 50%) of all *BRCA* mutation-positive families in this population group (van der Merwe & van Rensburg, 2009).

A single *BRCA2* founder mutation has been identified for the Coloured and Xhosa populations from the Western Cape. Although the mutation has been reported before to the BIC as 5999del4 detected in the Netherlands (c.5771_5774delTTCA,p.Ile1924_Gln1925fs), it is the first SA recurrent mutation detected in a non-Afrikaner population (van der Merwe *et al.*, 2012). Of a total of 105 Coloured and 16 Xhosa BC patients studied, 3.8% of the Coloured patients and 25% of the Xhosa patients harbored this mutation. Haplotype analysis indicated two distinct origins for the Netherlands and SA mutations (van der Merwe *et al*., 2012). The identification of these founder mutations made diagnostic testing for *BRCA* possible within certain SA population groups. Despite disease-causing single base changes distributed within the coding regions of both *BRCA1* and *BRCA2*, large rearrangements play an important role in the predisposition to BC. Large rearrangements in BC genes are the direct manifestation of repetitive *Alu* elements distributed unevenly in BC genes.

1.4.3 Larger genomic rearrangements within *BRCA1* and *BRCA2*

Although the familial BC genes *BRCA1* and *BRCA2* were identified more than a decade ago, it was not until 1997 that the first large genomic rearrangements (LGRs) present within *BRCA1* were reported (Puget *et al.*, 1997). The delay was in part attributed to the use of PCR-based techniques for the detection of point mutations and small insertions/deletions. These LGRs could not be detected simultaneously using these techniques, as they are not quantitative (Mazoyer, 2005).

The genomic regions of both *BRCA1* and *BRCA2* contain very high densities of repetitive DNA elements that have the potential to contribute to genomic instability (Fig. 1.7). Complete sequencing of *BRCA1* revealed a very high density of *Alu* sequences present within the gene (Smith *et al.*, 1996). *Alu* elements or sequences belong to a class of retroposons termed short interspersed elements (SINEs). Characteristics of these SINEs include a length of about 100–300 bp, they are commonly found in introns, in the 3’ untranslated regions (UTR) of genes and in in-

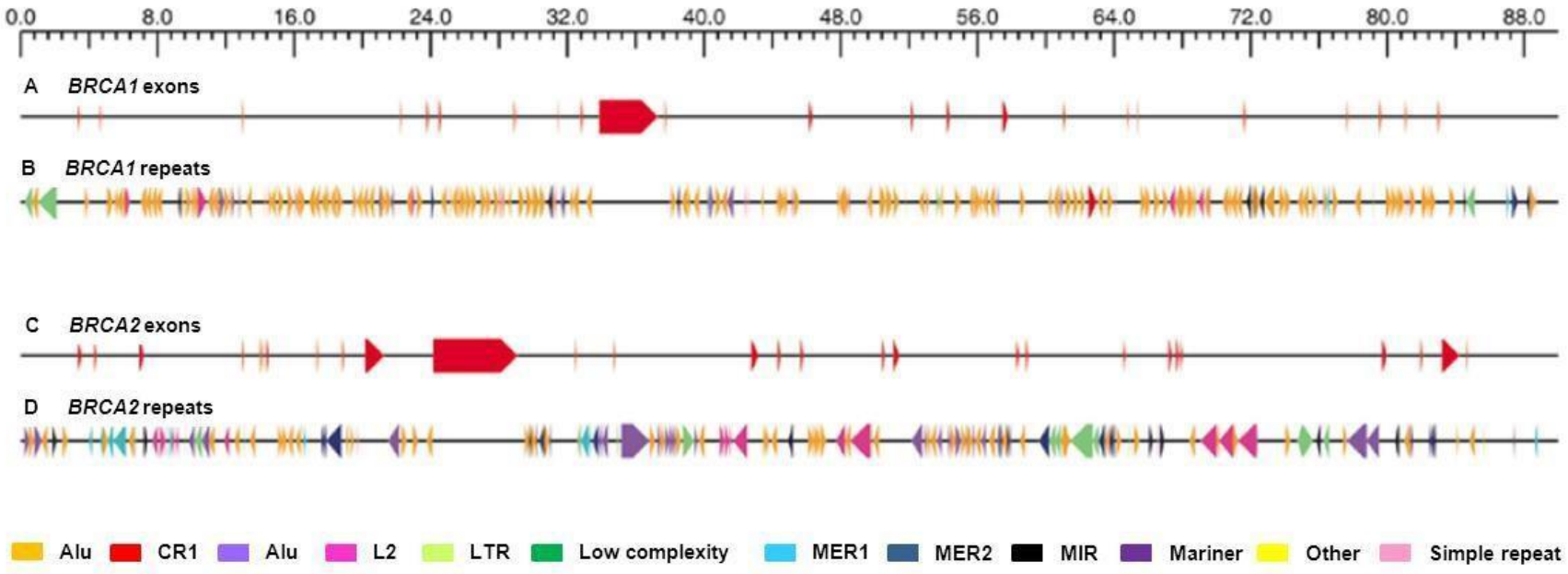

Figure 1.7 Repetitive DNA elements in the *BRCA1* and *BRCA2* genes. **A**. The genomic structure of *BRCA1*. The gene spans 84 kb of sequence and includes an unusually large central exon 11. **B**. Distribution of repetitive elements in *BRCA1*. *Alu* elements are highlighted in orange and comprise 42% of the gene. All repetitive elements comprise 47% of the gene. **C**. The genomic structure of *BRCA2*. The gene spans 86 kb of sequence and also includes an unusually large central exon 11. **D**. Distribution of repetitive elements in *BRCA2*. *Alu* elements are highlighted in orange and comprise 20% of the gene. The total percentage of repetitive elements is 48% of the gene. DrawMap by T Smith (unpublished data) was used by Welsch & King (2001) to create this figure.

tergenic genomic regions (Deininger & Batzer, 1993). These *Alu* repeats provide hotspots for unequal homologous recombination (Deininger & Batzer, 1999).

For *BRCA1*, there are 42% of *Alu* sequences and 5% of *non-Alu* repeats (Korenberg & Rykowski, 1988; Batzer *et al.*, 2002) (Fig. 1.7). *BRCA2*, on the other hand, comprises 47% of repetitive DNA of which only 20% consist of *Alu* sequences (Fig. 1.7). Although the two familial BC genes exhibit a high percentage of repetitive DNA elements, genes containing such a high density of repetitive DNA are rare. *Alu*-dense regions of the genome are associated with a high density of genes and localize predominantly to R bands of metaphase chromosomes, which are involved in homologous and non-homologous chromosomal exchange (Unger *et al*., 2000; Welsch & King, 2001; Belogianni *et al*., 2004; Bunyan *et al*., 2004; Zhang *et al*., 2011).

The presence of these *Alu* repeats provides the most common mechanism for the creation of LGRs observed in *BRCA1/2*. This mechanism entails *Alu*-mediated unequal homologous recombination, followed by non-homologous events such as *Alu*/non-*Alu* or non-*Alu*/non-*Alu*, and a recombination event between *BRCA1* and the *BRCA1* pseudogene (Sluiter & van Rensburg, 2011). These non-homologous recombination events frequently result in deletions and short insertions at the site of the deletion (Hastings *et al.*, 2009). This mechanism occurs due to the high number of *Alu* repeats (41.5%) present within *BRCA1*. This percentage is 4-fold higher than that observed for the human genome and is 2-fold that observed for *BRCA2* (Ewald *et al.*, 2009; Zhang *et al.*, 2011). Due to the presence of such a high percentage of *Alu* repeats in *BRCA1*, the prevalence of LGRs ranges from ~ 6 – 27% of all mutations detected in this gene (Petrij-Bosch *et al*., 1997; Lahti-Domenici *et al*., 2001; Hogervorst *et al.*, 2003; Laurila *et al*., 2005). In stark contrast, LGRs in *BRCA2* play a minimal role in BC (Sluiter & van Rensburg, 2011).

The contributions of LGRs to BC risk varies between the different populations, ranging from 0 – 27% (Hogervorst *et al.*, 2003; Pietshmann *et al.*, 2005; Moisan *et al.*, 2006). In some countries, it represents a founder mutation (such as the Netherlands) whereas in others LGRs are totally absent (Sluiter & van Rensburg, 2011). An example of the absence of LGRs is within the Ashkenazi Jewish

population (Stadler *et al.*, 2010). This study screened for the presence of LGRs in a total of 108 breast and OVC patients that previously tested negative for small mutations. Although these patients had an average pre-test mutation probability of 24.7% for the presence of LGRs, none were found (Stadler *et al.*, 2010). The most striking *BRCA1* founder mutations identified thus far internationally, include the 5.1 kb deletion of exon 22, the 3.8 kb deletion of exon 13 and the 6 kb insertion of exon 13 that is common amongst Europeans (The *BRCA1* exon 13 duplication screening group, 2000).

The largest study ever performed regarding the search for LGRs within the BC genes was by the privately owned Myriad Genetic Laboratories situated in the United States of America (Judkins *et al.*, 2012). As Myriad was the referral centre for the entire continent, they had LGRs data for a total of 48,456 BC patients gathered over a five year period (2007 – 2011). They made use of an in-house probe design assay (BART - The *BRCAnalysis* Rearrangement Test) that is similar to the Multiplex ligation-dependent probe amplification (MLPA) probemix introduced by MRC Holland.

A total number of 81 different LGRs in *BRCA1* and 27 in *BRCA2* (Fig. 1.8) were observed (Judkins *et al.*, 2012). These LGRs varied from the deletion or insertion of a single exon to whole-gene deletions of either *BRCA1* or *BRCA2* (Fig. 1.8). The group detected a total of 108 LGRs within this subset of patients. Many of the LGRs were detected for multiple patients (Fig. 1.8 A). Apart from the recurrent mutations, the majority of LGRs were detected in fewer than 10 patients, with many only occurring once (Judkins *et al*., 2012). The majority (≥90%) of LGRs were detected for *BRCA1* and was linked to the abundance of *Alu*- repeats present within the gene (Judkins *et al*., 2012) (Fig. 1.7). This finding is consistent with reports by various authors as tabulated within Reeves *et al*. (2006) as the majority of LGRs detected are *Alu* mediated (Fig. 1.7). The *Alu* repeats create “hotspots” ideal for unequal homologous recombination, which results in LGRs.

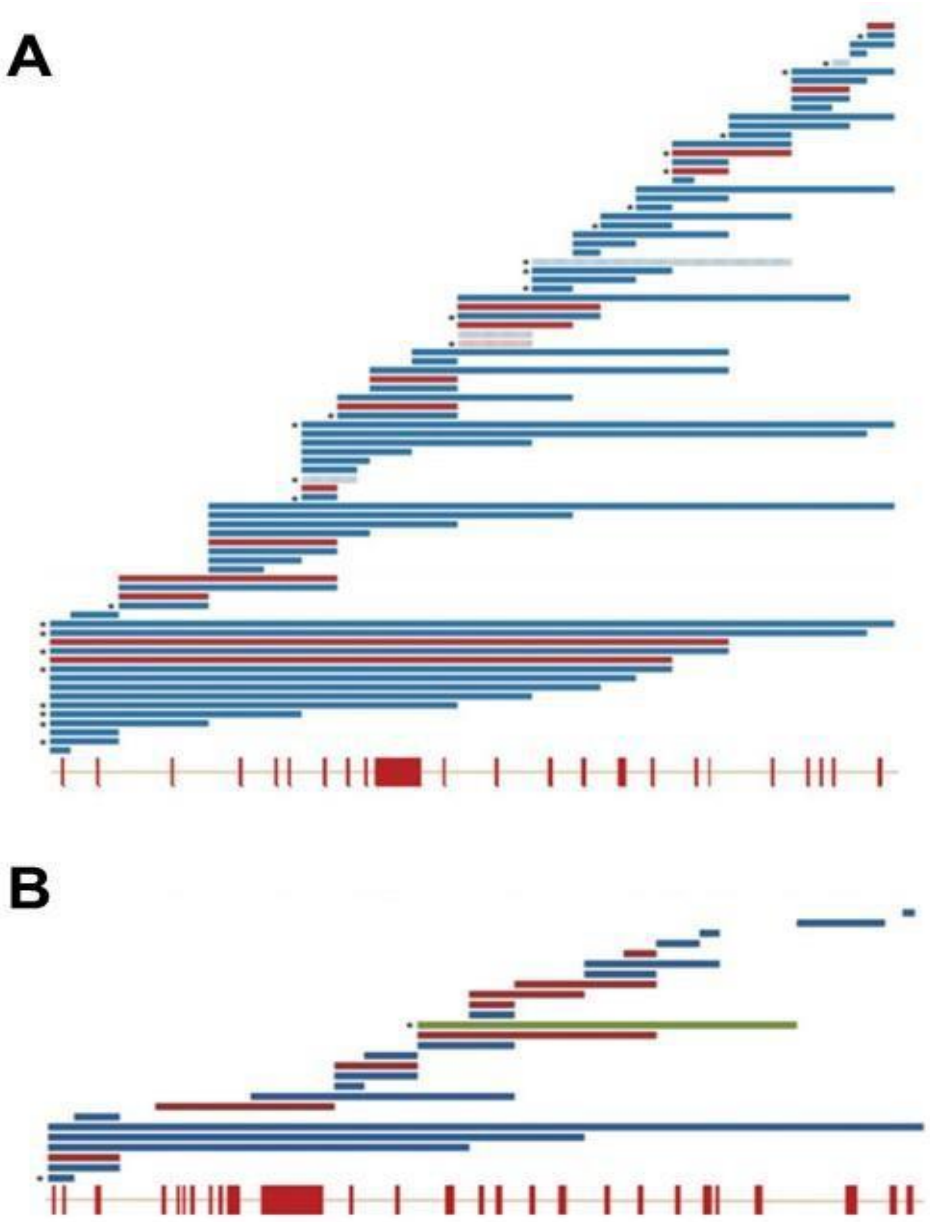

Figure 1.8 The spectrum of LGRs present within BC patients referred to Myriad Laboratories for analysis. Both genes start at exon 1, although *BRCA1* has no exon 4. Blue bars represent deletions, red bars indicate duplications, and the green bar represents a documented triplication. Rearrangements are indicated from the midpoint between affected exons for this schematic, therefore no actual breakpoint locations are implied. An asterisk (*) denotes rearrangements that were observed 5 or more times in this time period. **A**. LGRs detected for *BRCA1*. Five recurrent *BRCA1* rearrangements are indicated with hashed bars. **B**. LGRs detected for *BRCA2* (adapted from Judkins *et al.*, 2012).

1.4.3.1 Large genomic rearrangements present in Africa

In populations of African ancestry, the prevalence of LGRs within *BRCA1* and *BRCA2* genes has remained understudied. However, there have been attempts to screen LGRs in various African population groups by numerous studies. Reeves *et al.* (2006) performed MLPA for 56 Caucasian patients representing SA, but for *BRCA1* only. A single patient tested positive for an exon 13 deletion within that study (occurrence of 1.7%). Sluiter and van Rensburg (2011) completed this initial study by screening 52 breast/ovarian cancer families for the presence of LGRs, of which 36 represented Afrikaners families. They aimed to determine the impact of *BRCA1* and *BRCA2* LGRs in SA. They detected a single novel deletion, namely *BRCA1* exons 23 and 24 (NG_005905.2:g.169527_1805279del) on an SA family with Greek ancestry.

Another SA study was performed by Francies *et al.* (2015), screening 108 BC patients for the presence of LGRs. The subset included Black, Coloured, Caucasian, and Indian patients. Only a single Caucasian patient tested positive for a LGR (occurrence 0.9%), in which *BRCA1* exons 1a and 2 were deleted. These findings suggest that *BRCA* rearrangements in SA contributed about 3% to the *BRCA1* mutation burden (Sluiter & van Rensburg, 2011). This is a low rate in comparison to the other populations regarding the presence of LGRs.

In 2010, a study was published by Zheng *et al.* (2010) evaluating 352 Nigerian patients for the presence of LGRs also within *BRCA1*. Only a single patient tested positive for the presence of a deletion of exon 21 (c. 5277+480_5332+672del) (0.3% occurrence). The authors suggested that *BRCA1* genomic rearrangements exist but do not contribute significantly to *BRCA1* associated risk in the Nigerian population.

MLPA data also exist for Algeria (Cherbal *et al.*, 2010). The authors studied 86 high risk female BC patients who had an extensive family history of the disease. Both *BRCA1* and *BRCA2* were analysed using MLPA. Two recurrent mutations were identified, namely a novel single deletion of *BRCA1* exon 2 and *BRCA1* exon 8. These two mutations occurred in 8 of the 70 families (occurrence of 11.4%).

The data gathered by Myriad contributed indirectly to our knowledge regarding the African continent as their study included a total of 2714 African patients. In the Myriad study, Judkins *et al.* (2012) revealed that out of ~ 40000 patients, the highest overall LGRs detection rates were observed in patients of African and Latin American/Caribbean ancestry (29.4% and 31.2% respectively) (Judkins *et al.*, 2012). This is in stark contrast to the lowest positive rates (0.4% occurrence of 926 patients) observed for the Ashkenazi Jewish (the most extensively studied founder population), Native American and Asian ancestries. More LGRs were detected for *BRCA1* than *BRCA2* (Judkins *et al*., 2012). For *BRCA1*, the majority exhibited the *BRCA1* del exons 1-19 (detected for six patients), followed by *BRCA1* deletion (del) exon 8 (detected for five patients) and *BRCA1* duplication (dup) exons 18 - 19 (detected on 15 patients). These three mutations made up 60.5% of all the LGRs detected for the African population, with *BRCA1* dup exons 18-19 alone constituting 33% (Judkins *et al.*, 2012).

1.4.3.2 MLPA as detection method of large genomic rearrangements

Conventional DNA sequencing strategies, such as direct Sanger sequencing, are only capable of identifying small nucleotide sequence alterations. This can lead to an underestimation of the rate of mutations and therefore, a risk of a false-negative genetic report (Kwong *et al.,* 2015). Numerous methods have been utilized to detect LGRs, such as Southern blot and semi-quantitative multiplex PCR. These methods are high in cost, time consuming and labour intensive (Armour *et al.*, 2002).

The number of LGRs reported for *BRCA1/BRCA2* over the last few years has increased considerably. In 2005, only 29 *BRCA1* and 3 *BRCA2* LGRs were reported (Mazoyer *et al.,* 2005). However, in 2009, the numbers increased to 81 *BRCA1* and 17 *BRCA2* LGRs due to the wide use of MLPA. It is the most commonly used and effective method of detecting LGRs. It is a PCR-based, high-throughput technique that can determine the relative copy number of different DNA target sequences simultaneously.

MLPA is a multiplex-PCR method used for the detection of copy number variations in genomic sequences. The various SALSA® MLPA® probe mixes are designed and manufactured by MRC Holland (Netherlands). This technique has the ability to detect up to 50 different genomic DNA sequences in a single reaction and can distinguish sequences differing in only one nucleotide (Schouten *et al.,* 2002; Homig-Holzel & Savula, 2012). The method targets very small sequences (50 – 70 bp), which enables the technique to identify the frequent, single gene aberrations which are too small to be detected by FISH. The method can be performed using DNA, is of relatively low cost, and is technically uncomplicated (http://www.mrc-holland.com/, accessed December 19, 2015).

The method is based on the hybridization of two probe oligonucleotides immediately adjacent to the denatured target sequences (Fig. 1.9). Only when the two probe oligonucleotides are both hybridized to their adjacent targets, can they be ligated during the ligation reaction (Fig. 1.9). Probe oligonucleotides that are not ligated, only contain one primer sequence. As a consequence, they cannot be amplified exponentially and will therefore not generate a signal. The removal of unbound probes are therefore unnecessary and makes the MLPA method easy to perform. Once ligation has occurred, the ligated probes are amplified and the amplification products are separated by capillary electrophoresis (Fig. 1.9). The resulting amplification products of a SALSA® MLPA® probemix range between 1.3 and 4.80 kb in length. The final step is data analysis (Fig. 1.10).

A change in copy number of the MLPA probe-target sequences results in a lower or higher relative amount of the probe amplification product (Fig. 1.10). An MPLA reaction makes it possible to detect heterozygous deletions (Fig. 1.10 C) and amplifications by comparing the relative signal of the probe with the average relative signal of the same probe in the normal reference sample (Homig-Holzel & Savola 2012) (http://www.mrc-holland.com/, accessed December 19, 2015).

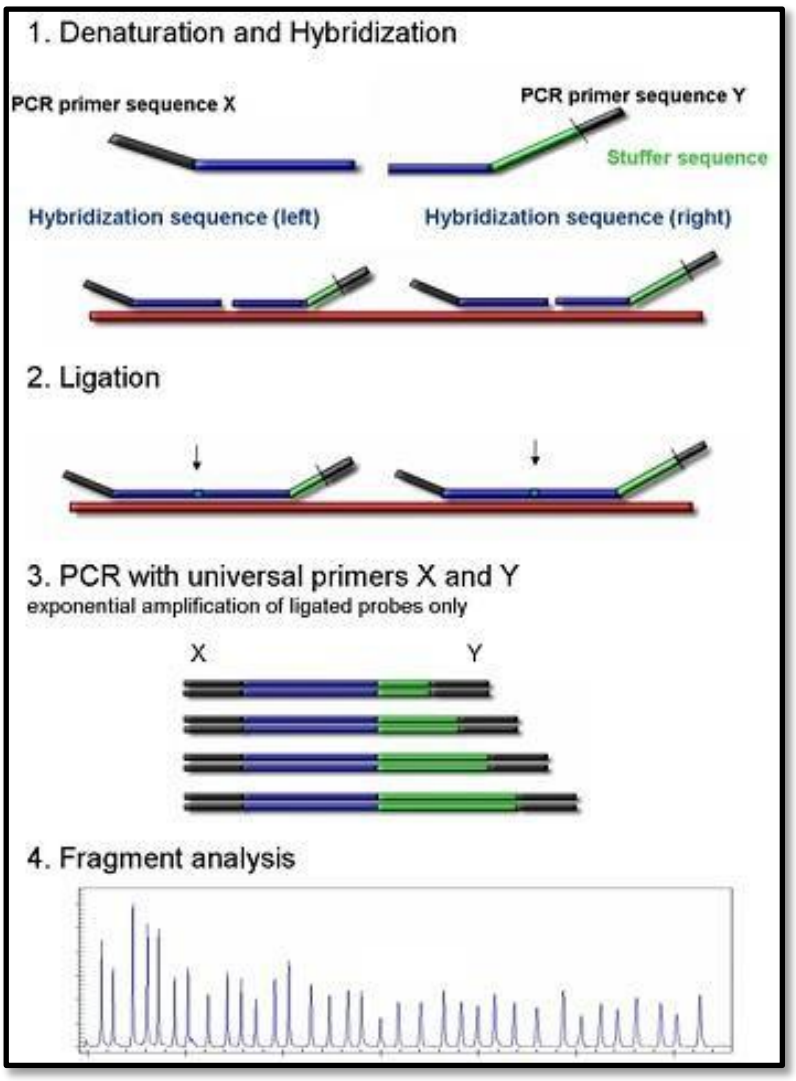

Figure 1.9 An illustration of MLPA reaction using two hemi probes for sample amplification. (http://www.mlpa.com/WebForms/WebFormMain.aspx?Tag accessed on 17 October 2015).

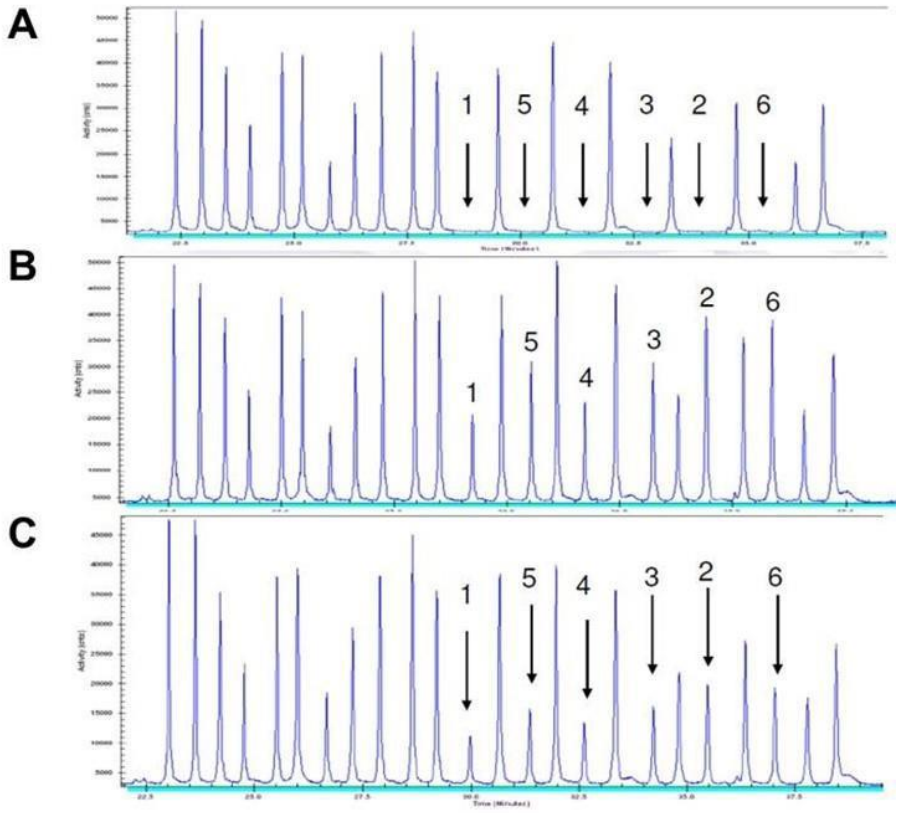

Figure 1.10 Identification of LGRs within the *ASPA* gene using MLPA. **A**. Homozygous deletion of exons 1 - 6 within the *ASPA* gene, as indicated by the arrows. **B**. Reference sample indicating normal peak heights for exons 1 – 6, as indicated by the numbers. **C**. Heterozygous deletion of exons 1 - 6 within the *ASPA* gene, as indicated by the arrows (Schouten *et al.*, 2002).

1.5 Objectives of the study

Currently, no comprehensive SA study exists that can shed light on the contribution that LGRs make towards the *BRCA* mutation spectrum within the SA population. Only by gathering evidence, we would be able to determine the frequency and presence if any, of these large genomic rearrangements within our various population groups. The aims of this study are:

- To implement and validate the MLPA technique for routine use within the diagnostic laboratories of the Division of Human Genetics and
- To screen breast cancer patients who have previously tested negative for the other *BRCA* mutations in patients representing specifically the Coloured, the Black and the Indian population for the presence of LGRs.

Chapter 2

Material and Methods

2.1 Patients

The Division of Human Genetics of the National Health Laboratory Services (NHLS) in Bloemfontein has been the referral laboratory for familial BC for more than a decade. Initially, only a limited number of patients were analysed, but as the public awareness increased regarding the impact of these genes, more comprehensive screening is requested. For this reason, a diagnostic comprehensive screening was implemented to screen every coding exon and the splice-site boundaries for mutations within *BRCA1* and *BRCA2*, using a variety of mutation detection techniques.

Breast cancer patients representing various population groups of SA (Caucasian, Black African, Coloured, and Indian) have thus been diagnostically screened for the presence of small pathogenic mutations such as deletions, duplications and single base changes. These patients were initially selected for diagnostic testing based on clinical criteria predicting a relatively high probability of carrying a *BRCA1/2* mutation and the presence of a positive family history. These diagnostic patients were counselled by genetic counsellors at the various referral centres prior to diagnostic testing. The counselling session was structured according to the discipline specific guidelines for familial BC, as stipulated by the SA Genetic Counsellors website (http://www.geneticcounselling.co.za). Once the patient understood the information and was given ample time to ask questions, they gave informed consent (Appendix A) and testing proceeded.

From these patients, individuals were retrospectively selected based on various criteria. These included the following: patient must be diagnosed with invasive or *in-situ* BC, tested negative for the presence of smaller pathogenic *BRCA1* or *BRCA2*

mutations, represent one of three SA population groups (self-identified as either Black African, Coloured or Indian), has a minimum of one other family member affected with the disease or an early age at onset (under the age of 45).

A total of 129 patients were selected and included 58 Black African (Table 2.1), 29 Coloured (Table 2.2) and 42 Indian patients (Table 2.3). These patients were received from Inkosi Albert Luthuli Central Hospital in Kwazulu-Natal and Tygerberg Hospital in Cape Town. The Division of Human Genetics in Bloemfontein did not have any direct contact with these patients.

2.2 Ethics

This project was approved by the Evaluation Committee of the Faculty of Health Sciences, School of Medicine in Bloemfontein where after ethical approval was obtained from the Ethics Committee of the Faculty of Health Sciences of the University of the Free State in Bloemfontein for each of the population groups (ETOVS 65/08, Appendix B; ECUFS 107/2014, Appendix C; ECUFS 108/14, Appendix D). Permission was also requested and received from the Inkosi Albert Luthuli Central Hospital, the Province of KwaZulu-Natal and Tygerberg Hospital to make use of their patients (Appendix E, F & G). Permission was also obtained from the Business Manager of NHLS in Bloemfontein for the use of the laboratory space and apparatus of the Molecular Genetics laboratory (Appendix H).

2.3 Methodology

2.3.1 DNA extraction

After obtaining informed consent, 10 – 20 ml of blood was collected in ethylenediaminetetraacetic acid (EDTA) tubes by a qualified health worker at the referring institution. Once the blood was received in the Molecular Genetics laboratory in Bloemfontein, it was transferred to two labelled Nunc tubes and stored at -20°C. DNA was extracted according to the salting out method (Adapted from Miller *et al*., 1988). The blood sample was thawed and cells ruptured using 45 ml

Table 2.1 Compilation of Black patients included in this study, indicating the allocated laboratory number, *BRCA1*, and *BRCA2* non-pathogenic variants detected, the extent of the family history (if any) and the patient's age at diagnosis.

Patient number	Laboratory number	*BRCA* variants detected	Cancers present within the family	Age at diagnosis
1	CAM2225	*BRCA2* c.9058A>T,p.Ile2944Phe	1 BC (dx=27)	27
2	CAM2485	*BRCA2* c.8487+47C>T *BRCA2* c.10023C>T,p.Asn3341=	1 BC (dx=31)	31
3	CAM2514	*BRCA1* c.4115G>A,p.Cys1372Tyr	1 BC (dx=41)	41
4	CAM2538	None	1 BC, 1 throat (dx ≥36)	36
5	CAM2564	*BRCA1* c.4485-64C>G *BRCA2* c.8755-66T>C	1 BC (dx=31)	31
6	CAM2571	None	1 BC (dx=37)	37
7	CAM2577	*BRCA1* c.302-41T>C *BRCA1* c.4843A>G,p.Ser1613Gly	1 BC (dx=55)	55
8	CAM2612	None	1 BC (dx=27)	27
9	CAM2697	None	6 BC (dx≥30)	65
10	CAM2702	None	1 BC (dx=48)	48
11	CAM2714	None	3 BC (dx≥30)	58
12	CAM2729	None	1 BC, 1 prostate ca	45
13	CAM2734	None	1 BC (dx=38)	38
14	CAM2741	*BRCA1* c.212G>A,p.R71K *BRCA1* c.4837A>G,p.S1613G	3 BC (all dx 20)	47
15	CAM2743	*BRCA2* c.681+56C>T *BRCA2* c.8755-75T>C	3 BC (dx≤50)	30
16	CAM2744	*BRCA2* c.681+56C>T *BRCA2* c.1114A>C,p.Asn372His *BRCA2* c.8488-31C>A *BRCA2* c.8830A>T,p.Ile2944Phe	1 BC (dx≥50)	71
17	CAM2753	None	3 BC, 1 throat ca, 1 OVC (all dx≥50)	58
18	CAM2754	*BRCA1* c.21C>T,p.Arg7= *BRCA2* c.8755-66T>C	2 BC, 1 womb ca (all dx≥40)	47
19	CAM2756	*BRCA1* c.4987-20A>G	1 BC (dx=28)	28
20	CAM2759	None	1 BC (dx=35)	35
21	CAM2766	*BRCA1* c.4837A>G,p.Ser1613Gly	1 BC (dx=29)	29
22	CAM2768	None	1 BC (dx=38)	38
23	CAM2778	None	2 BC (dx≥30)	48
24	CAM2804	*BRCA2* c.7242A>G,p.Ser2414=	1 BC (dx=47)	47
25	CAM2815	*BRCA2* c.1114A>C,p.N372H	1 BC, 1 trachea ca (dx≥20	27
26	CAM2816	*BRCA2* c.100023C>T,p.Glu3344=	4 BC (all dx≥20)	24

27	CAM2829	*BRCA2* c.517-4C>G *BRCA2* c.7806-14T>C *BRCA2* c.8092C>A,pAla2698Thr	1 BC (dx=25)	25
28	CAM2846	*BRCA1* c.681+56C>T	1 BC, 1 OVC (dx≤ 46)	46
29	CAM2847	*BRCA2* c.10234A>G,p.I3412V	2 BC, 1 OVC (all dx≥40)	46
30	CAM2851	*BRCA1* c.4113G>A,p.Gly1317=	3 BC (dx≥30)	34
31	CAM2873	*BRCA1* c.4308T>C,p.Ser1436= *BRCA1* c.4837A>G,p.Ser1613Gly	1 BC (dx=52)	52
32	CAM2889	*BRCA2* c.381+56C>T *BRCA2* c.8487+47C>T *BRCA2* c.8755-75T>C	3 BC (dx≥57)	57
33	CAM2910	*BRCA1* c.4113G>A,p.Gly1371= *BRCA1* c.4185+12_4185+13delGT	1 BC (dx=43)	43
34	CAM2912	*BRCA1* c.670+59C>T *BRCA1* c.4837A>G,p.Ser1613Gly *BRCA2* c.7017G>C,p.Lys2339Asn *BRCA2* c.7319A>G,p.His2440Arg *BRCA2* c.7397C>T,p.Ala2466Val *BRCA2* c.8755-66T>C *BRCA2* c.9730G>A,p.Val3244Ile	2 BC (dx≥30)	38
35	CAM2920	*BRCA2* c.681+56C>T *BRCA2* c.7806-14T>C	1 BC (dx=30)	30
36	CAM2928	*BRCA2* c.-26G>A *BRCA2* c.516+21delA *BRCA2* c.681+56C>T	1 BC, 1 OVC, 1 uterus (dx≥30)	36
37	CAM2932	*BRCA2* c.6842-73T>A	1 BC (dx=31)	31
38	CAM2946	*BRCA2* c.4837A>G,p.S1613G	1 BC (dx=28)	28
39	CAM2947	*BRCA2* c.-26G>A	1 BC (dx=45)	45
40	CAM2948	*BRCA1* c.4308T>C,p.Ser1436= *BRCA1* c.4837A>G,p.Ser1613Gly *BRCA2* c.681+56C>T *BRCA2* c.7806-14T>C *BRCA2* c.8755-66T>C	1 BC (dx=35)	35
41	CAM2949	*BRCA2* c.7397C>T,p.Ala2466Val	1 BC, 1 stomach ca (dx≥30)	35
42	CAM2950	*BRCA2* c.7242A>G,p.Ser2414= *BRCA2* c.-26G>A *BRCA2* c.681+56C>T *BRCA2* c.8487+19A>G *BRCA2* c.8755-66T>C	1 brain ca (dx≤35)	35

43	CAM2958	*BRCA1* c.670+59C>T *BRCA1* c.4837A>G,p.Ser1613Gly *BRCA2* c.6842-73T>A *BRCA2* c.1023A>G,p.Ile3412Val	2 BC (dx≤60)	59
44	CAM2965	*BRCA1* c.4308T>C,p.Ser1436= *BRCA1* c.4837A>G,p.Ser1613Gly *BRCA2* c.1114A>C,p.Asn372His *BRCA2* c.7242A>G,p.Ser2414= *BRCA2* c.8755-66T>C *BRCA2* c.8830A>T,p.Ile2944Phe	2 BC (dx≥30)	34
45	CAM2967	*BRCA2* c.467A>G,p.D156G *BRCA2* c.1114A>C,p.N372H *BRCA2* c.8830A>T,p.I2944F	1 BC (dx=31)	31
46	CAM2971	None	1 BC (dx=36)	36
47	CAM2972	*BRCA1* c.4113G>A,p.Cys1372= *BRCA1* c.4185+12_4185+13delGT *BRCA2* c.681+56C>T *BRCA2* c.1114A>C,p.Asn372His *BRCA2* c.8755-66T>C *BRCA2* c.8830A>T,p.Ile2944Phe	1 BC (dx=37)	37
48	CAM2981	*BRCA2* c.681+56C>T *BRCA2* c.7242A>G,p.Ser2414= *BRCA2* c.8755-6 6T>C	3 BC (dx≥40)	40
49	CAM2982	*BRCA1* c.212+23T>A *BRCA1* c.301+8T>C *BRCA1* c.5406+68T>C *BRCA2* c.467A>G,p.Asp156Gly *BRCA2* c.8487+8G>T	1 BC (dx=35)	35
50	CAM2988	*BRCA1* c.4987-20A>G *BRCA2* c.-26G>A *BRCA2* c.681+56C>T *BRCA2* c.7242A>G,p.Ser2414= *BRCA2* c.8755-66T>C	1 BC (dx=43)	43
51	CAM2989	*BRCA2* c.-26G>A *BRCA2* c.7242A>G,p.Ser2414= *BRCA2* c.7397C>T,p.A2466V	1 BC (dx=45)	45
52	CAM2990	*BRCA1* c.4185+12_4185+13delGT *BRCA2* c.317-22C>T *BRCA2* c.7324C>G,p.Gly2441= *BRCA2* c.1023A>G,p.Ile3412Val	1 BC, 1 Leukaemia (dx≥30)	35

53	CAM2991	*BRCA2* c.681+56C>T *BRCA2* c.8406A>C,p.Val2820= *BRCA2* c.8755-66T>C *BRCA2* c.10234A>G,p.Ile3412Val	2 BC (dx≥35)	35
54	CAM2993	*BRCA1* c.4837A>G,p.Ser1613Gly *BRCA2* c.7242A>G,p.Ser2414= *BRCA2* c.7397C>T,p.Ala2466Val *BRCA2* c.8755-56T>C	1 BC, 1 cervix (dx≥30)	34
55	CAM2995	*BRCA1* c.4308-56T>C,perSer1436=	1 BC, 1 OVC (dx≤50)	27
56	CAM2996	*BRCA2* c.7017G>C, p.Lys2339Asn *BRCA2* c.7397C>T, p.A2466V *BRCA2* c.10234A>G,p.Ile3412Val	1 BC, 1 prostate ca (dx≥40)	52
57	CAM2997	*BRCA2* c.7017G>C, p.Lys2339Asn *BRCA2* c.7324C>G,p.Gly2441= *BRCA2* c.7397C>T, p.A2466V	1 BC (dx=38)	38
58	CAM3013	*BRCA2* c.-26G>A *BRCA2* c.681+56C>T *BRCA2* c.793+65delT *BRCA2* c.7806-14T>C *BRCA2* c.8487+47C>T	2 BC(dx≥33)	33

Table 2.2 Compilation of Coloured patients included in this study, indicating the allocated laboratory number, *BRCA1*, and *BRCA2* non-pathogenic variants detected, the extent of the family history (if any) and the patient's age at diagnosis.

Patient number	Laboratory number	*BRCA* variants detected	Cancers present within the family	Age at diagnosis
1	CAM1482	None	1 BC, 1 brain, 1 stomach, 1 lung (dx≥50)	37
2	CAM1497	None	2 BC (dx≥30)	36
3	CAM2267	*BRCA1* c.5406+T>C	1 BC (dx=54)	54
4	CAM2271	*BRCA2* c.7470A>G,p.Ser2414=	3 BC (dx≥40)	43
5	CAM2318	*BRCA1* c.4485-64C>G	1 BC (dx=32)	32
6	CAM2349	*BRCA1* c.114G>A,p.Lys38Asn	2 BC (dx≥40)	52
7	CAM2361	*BRCA1* c.4308T>C p.Ser1436= *BRCA2* c.8755-66T>C	6 BC, 5 stomach ca (dx≥30)	58
8	CAM2378	*BRCA1* c.548-58_548-delT	2 BC, 1 OVC (dx≥40)	68
9	CAM2496	*BRCA2* c.8182G>A,p.Val2728Leu *BRCA2* c.8755-66T>C	2 BC, 4 prostate (dx≤60)	26
10	CAM2535	None	2 BC (dx≤40)	28
11	CAM2642	None	1 BC (dx=33)	33
12	CAM2703	*BRCA2* c.-26G>A *BRCA2* c.7242A>G,p.Ser2414=	3 BC, 1 thyroid (dx≥30)	62
13	CAM2717	None	3 BC (all dx≥35)	39
14	CAM2731	*BRCA1* c.4308T>C,p.Ser1436= *BRCA1* c.4837A>G,p.Ser1613Gly *BRCA1* c.4956G>A,p.M1652I *BRCA2* c.-26G>A *BRCA2* c.7017G>C,p.Lys2339Asn	1 BC (dx=26)	26
15	CAM2755	*BRCA2* c.425+67A>C *BRCA2* c.865A>C,p.Asn289His *BRCA2* c.1114A>C,p.N372H *BRCA2* c.8755-66	3 BC, 1 colon, 1 prostate (dx≤60)	45
16	CAM2761	*BRCA1* c.4113G>A,p.Gly1371=	1 BC, 2 lung ca, 1 cervix ca (all dx≥30)	39
17	CAM2767	*BRCA2* c.1114A>C,p.Asn372His *BRCA2* c.7806-14T>C *BRCA2* c.8755-66T>C	4 BC (dx≤60)	53
18	CAM2782	*BRCA1* c.442-34C>T	5 BC, 5 prostate (all dx≥50)	51
19	CAM2785	*BRCA1* c.425+67A>C *BRCA1* c.426-89C>T *BRCA1* c.681+56C>T	1 BC, 2 OVC (all dx≥40)	50

20	CAM2811	*BRCA1* c.442-34C>T *BRCA1* c.4308T>C,p.Ser1436=	2 BC, 1 OVC (all dx≥30)	57
21	CAM2854	*BRCA2* c.-26G>A *BRCA2* c.1114A>C,p.N372H *BRCA2* c.7242A>G,p.Ser2414	3 BC (dx≥30)	30
22	CAM2860	*BRCA1* c.442-36C>T *BRCA1* c.4837A>G,p.Ser1613Gly	3 BC (dx≥39)	39
23	CAM2861	*BRCA2* c.681+56C>T *BRCA2* c.8487+47C>T *BRCA2* c.8755-66T>C	2 BC (dx≥28)	28
24	CAM2882	*BRCA1* c.4308T>C,p.Ser1436= *BRCA1* c.4837A>G,p.Ser1613Gly *BRCA2* c.1114A>C,p.N372H	3 BC (dx≥40)	64
25	CAM2883	None	3 BC (dx≥26)	26
26	CAM2933	*BRCA1* c.442-34C>T *BRCA2* c.-26G>A *BRCA2* c.1114A>C,p.Asn372His	4 BC, 1 stomach (dx≥40)	43
27	CAM2943	*BRCA2* c.7017G>C,p.Lys2339ASn *BRCA2* c.7242A>G,p.Ser2414= *BRCA2* c.9730G>A,p.Val3244Ile *BRCA2* c.10023C>T,p.Asp3344=	2 BC (dx≤40)	33
28	CAM2983	*BRCA2* c.681+56C>T *BRCA2* c.8755-66T>C *BRCA2* c.1023A>G,p.Ile3412Val	1 BC (dx=31)	31
29	CAM2987	*BRCA1* c.212+66A>G *BRCA1* c.4308T>C,p.Ser1436= *BRCA2* c.5406+8T>C *BRCA2* c.-11C>T *BRCA2* c.-26G>A *BRCA2* c.681+56C>T *BRCA2* c.8242A>G,p.Ser2414= *BRCA2* c.8755-66T>C	2 BC (dx≥45)	45

Table 2.3 Compilation of SA Indian patients included in this study, indicating the allocated laboratory number, *BRCA1*, and *BRCA2* non-pathogenic variants detected, the extent of the family history (if any) and the patient's age at diagnosis.

Patient number	Laboratory number	*BRCA* variants detected	Cancers present within the family	Age at diagnosis
1	CAM2113	*BRCA1* IVS7-34T>C	1 BC, 2 stomach ca (dx≤50)	29
2	CAM2184	None	2 BC (dx≥30)	30
3	CAM2247	*BRCA1* c.4873A>G,p.Ser1613Gly	3 BC (dx≥20)	21
4	CAM2293	None	3 BC, 1 lung ca(dx≥50)	56
5	CAM2333	*BRCA1* c.4308T>C,p.Ser1436=	1 BC (dx=30)	30
6	CAM2335	*BRCA1* c.4308T>C, p.Ser1436= het	3 BC,1 throat ca (dx≥35)	35
7	CAM2336	None	2 BC (dx≥35)	46
8	CAM2338	*BRCA2* c.8755-66T>C	2 BC (dx≥30) 2 OVC (all dx≥35)	35
9	CAM2339	*BRCA1* c.442-34C>T	2 BC (all dx≥35)	35
10	CAM2473	*BRCA1* c.4308T>C,p.Ser1436= *BRCA1* c.4837A>G,p.Ser1613Gly	2 BC, 2 BC/OVC (dx≤60)	60
11	CAM2481	*BRCA2* c.-26 G>A *BRCA2* c.7242A>G,p.S2414= *BRCA2* c.8755-75T>C	2 BC,1 lung ca (dx≤40)	37
12	CAM2513	None	2 BC (dx≥30)	31
13	CAM2524	*BRCA1* c.4873A>G,p.Ser1613Gly *BRCA2* c.5152+73T>C	3 BC (dx≥28)	28
14	CAM2604	*BRCA2* c.-26G>A *BRCA2* c.8755-66TC	2 BC (dx≥34)	34
15	CAM2644	*BRCA1* c.4308T>C,p.Ser1436= *BRCA1* c.4873A>G,p.Ser1613Gly	3 BC (dx≥21)	21
16	CAM2645	*BRCA2* c.425+67A>C *BRCA2* c.865A>C.p.Asn289His	2 BC (dx≥30)	35
17	CAM2647	*BRCA1* c.4308T>c,p.Ser1436= *BRCA1* c.4837A>G,p.Ser1613Gly	2 BC (dx≥39)	39
18	CAM2708	*BRCA1* c.442-34C>T	4 BC (dx≥40)	48
19	CAM2710	*BRCA1* c.4308T>C,p.Ser1436=	1 BC, 1Stomach ca, 1 lung ca (all dx≥40)	46
20	CAM2737	*BRCA2* c.1114A>C,p.Asn372His	1 BC	30
21	CAM2742	*BRCA2* c.1114A>C,p.N372H	3 BC (dx≤68)	68
22	CAM2750	*BRCA1* c. 591C>T,p.Cys197=	5 BC (all dx≥30)	32
23	CAM2758	*BRCA1* c.-26G>A	3 BC, 1 uterus ca (all dx≥20)	20
24	CAM2780	*BRCA2* c.7242A>G,p.Ser2414=	3 BC, 1 throat (all dx≥40)	66

25	CAM2806	*BRCA1* c.681+56C>T	2 BC, 1 liver ca, 1 lung ca, 1 stomach ca (all dx≥40	57
26	CAM2820	*BRCA1* c.4308T>C,p.Ser1436= *BRCA1* c.4837A>G,p.Ser1613Gly	3 BC (all dx≥50)	67
27	CAM2821	*BRCA1* c.-26G>A *BRCA1* c.681+56C>T	3 BC, 1 stomach ca (all dx≥ 20)	48
28	CAM2823	None	3 BC (dx≥32)	32
29	CAM2826	*BRCA1* c.4308T>C,p.Ser1436=	3 BC (all dx≥30)	67
30	CAM2841	None	5 BC (all dx≥19)	19
31	CAM2850	*BRCA1* c.442-34C>T	1 BC (dx=30)	30
32	CAM2872			
33	CAM2892	*BRCA2* c.425+67A>C *BRCA2* c.426-89T>C *BRCA2* c.865A>C p.Asn289His	1 BC, I OVC (dx≤30)	28
34	CAM2897	*BRCA1* c.4308T>C,p.Ser1436= *BRCA1* c.4387A>G,p.Ser1613Gly *BRCA2* c.-26G>A *BRCA2* c.1114A>C,p.N372H	2 BC, 1 leukemia (dx≤69)	69
35	CAM2911	*BRCA1* c.591C>T,p.Cys197= *BRCA1* c.4987-20A>G	7 BC (dx ≥30)	42
36	CAM2915	*BRCA1* c.442-34C>T	3 BC, 1 prostate ca (dx≥40)	61
37	CAM2916	*BRCA2* c.1114A>C p.Asn372His	1 BC, 1 OVC (dx≥40)	62
38	CAM2917	*BRCA1* c.442-34C>T *BRCA1* c.4308T>C,p.Ser1436=	3 BC (dx≤50)	35
39	CAM2954	*BRCA2* c.1114A>C,p.Asn372His	3 BC (dx≥58)	58
40	CAM2960	*BRCA1* c.4308T>C,p.Ser1436= *BRCA1* c.4837A>G,p.ser1613Gly *BRCA2* c.1114A>C,p.Asn372His	2 BC (dx≤68)	68
41	CAM2962	*BRCA1* c.4308T>C,p.Ser1436= *BRCA1* c.4837A>G,p.Ser1613Gly *BRCA2* c.1114A>C,p.Asn372His	3 BC (dx≤75)	75
42	CAM2963	None	4 BC, 1 Uterus (dx≥45)	55

cold lysis buffer [0.3 M sucrose,10 mM 2-amino-2-(hydroxymethyl)-1,3-propanediol (Tris) pH 7.8, 5 mM $MgCl_2$, 1% (v/v) t-octylphenoxypolyethoxyethanol (Trixton X-100)]. The solution was centrifuged for 20 min at 4 000 *g* at 4°C, where after the pellet was resuspended in 1x SET buffer (10 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM EDTA) containing 10 $\mu g.\mu l^{-1}$ proteinase K and 1% (w/v) sodium dodecyl sulphate (SDS). The solution was incubated overnight at 37°C.

A volume of 1.4 ml saturated NaCl was added the following morning, where after the tubes were vigorously shaken for 15 sec and centrifuged for 15 min at 4 000 *g* at 4°C. This step was repeated twice. After centrifugation, the supernatant was transferred to a clean tube and the DNA precipitated using 2 volumes 100% (v/v) ethanol. The solution was gently mixed and incubated for 10 min at -20°C. The precipitated DNA was then scooped from the solution, transferred to an Eppendorf tube and washed with 70% (v/v) ethanol. This step was repeated twice, where after the DNA was air dried and dissolved in $T._1E$ buffer (10 mM Tris-HCl pH 7.6, 0.1 mM EDTA).

If no DNA was visible after precipitation, the tubes were incubated at -20°C overnight. The tubes were then centrifuged for 30 min at 4 000 *g* (4°C) to pellet the DNA, where after the pellet was washed with 70% (v/v) ethanol, air dried and dissolved in $T._1E$ buffer (Miller *et al*., 1988).

2.3.2 DNA concentration determination

The quantity and purity of the extracted DNA were determined by making use of the NanoDrop®ND-1000 Spectrophotometer (v3.01, NanoDrop Technologies) according to the manufacturer's protocol. The quantity of the sample was expressed as $ng.\mu l^{-1}$. The purity of the extracted DNA was determined by reading two values at the following wavelengths ($A_{260/280}$ and $A_{230/260}$). These ratios were used to determine the presence of contaminants such as proteins or other contaminants which absorb at 230nm and 280nm (Technical Bulletin, NanoDrop – www.nanodrop.com/.../T042-NanoDrop-Spectrophotometers-Nucleic-Acid).

2.3.3 Sample preparation

The DNA stock solution was diluted using $T._1E$ buffer containing 5 - 10 mM Tris buffer (pH 8.0). The DNA dilutions needed for MLPA were meticulously diluted the day before analysis and placed overnight at 37°C to obtain a homogenous final concentration of 50ng.μl^{-1} in a 5 µl volume. Three reference samples were used during every run to evaluate the quality and the reproducibility of each probe within the run. One of the reference samples included a male individual and was used to correlate the sex of the individual with that of the X and Y probes present within the probe mix. A no DNA control was also included within each run to check for contamination of any of the MLPA reagents. For the final validation process, DNA samples for which smaller and larger *BRCA* mutations have been confirmed by an international accredited laboratory were used. These samples acted as positive controls.

2.3.4 MLPA method

MLPA was performed according to the MLPA® DNA Protocol version MDP-005 as last revised on 22 September 2014. The technique utilized the following SALSA® MLPA® kits for *BRCA1*, namely the older probemix P002-C1 (version 24 issued 01 March 2011, Table 2.4), the current probemix P002-D1 (version 36 issued 25 February 2015, Table 2.5) and probemix P087-C1 (version 23 issued 19 June 2014, Table 2.6). The latter served as a confirmation kit in the case positive results were obtained with either probemix P002-C1 or P002-D1. For *BRCA2*, the current SALSA® MLPA® probemix P045-B3 *BRCA2/CHEK2* (version 33 issued 09 September 2015, Table 2.7) and the confirmation kit P077-A3 (version 6 issued 17 November 2014, Table 2.8) were used (MRC-Holland, Amsterdam, the Netherlands), according to the manufacturer's instructions.

SALSA® MLPA® probemix P002-D1 included probes for each of the 24 exons of *BRCA1* (Table 2.5), whereas SALSA® MLPA® probemix P045-B3 included probes for the 27 exons of *BRCA2*, together with exon 9 of *CHEK2* (Table 2.7) (version 23 issued 19 July 2014, Table 2.8). A few minor adaptions were tested during the optimization process. The method can be divided into four phases, namely hybridization of the probes to the target region, ligation of the probes, the amplification of the reaction and fragment analysis.

Table 2.4 The *BRCA1* probes included in the SALSA® MLPA® P002-C1 probemix. Indicated are the length of each of the probes, SALSA MLPA probe number, exon number, the ligation site with the partial ligation surrounding sequences and the distance from the next probe (as described in COA Version 24, issued on 01/03/2011).

Length (nt)	SALSA MLPA probe	BRCA1 exon	Ligation site	Partial Sequence	Distance to next probe
		Start codon	233-235		
148	00763-L00268	Exon 1A	192-191 reverse	AGCAGAGGGTGA-AGGCCTCCTGAG	0.2 kb
157	00764-L00269	Intron 1	208 nt after exon 1A	AGGGGGCACTGA-GTGTCCGTGGGG	1.0 kb
166	00765-L00270	Exon 2	248-249	ATTTATCTGCTC-TTCGCGTTGAAG	8.3 kb
175	00826-L00341	Exon 3	335-336	TCAAGGAACCTG-TCTCCACAAAGT	9.3 kb
184	00767-L00272	Exon 5	393-394	ACTTCTCAACCA-GAAGAAAGGGCC	1.6 kb
208	00827-L00342	Exon 6	473-474	CGAGATTTAGTC-AACTTGTTGAAG	0.8 kb
216	00769-L00274	Exon 7	637-638	AACCGTGCCAAA-AGACTTCTACAG	4.3 kb
226	01004-L00569	Exon 8	718-719	CTTGGAACTGTG-AGAACTCTGAGG	2.6 kb
236	01005-L00581	Exon 9	813-814	CGTTAATAAGGC-AACTTATTGCAG	1.3 kb
244	00772-L00277	Exon 10	853-854	TTGTTACAAATC-ACCCCTCAAGGA	1.1 kb
266	00830-L00345	Exon 11	994-995	AAGCGTGCAGCT-GAGAGGCATCCA	2.8 kb
277	00774-L00279	Exon 11	3810-3811	TCCTAGCCCTTT-CACCCATACACA	1.0 kb
285	00775-L00280	Exon 12	4377-4378	CTCTGAAGACTG-CTCAGGGCTATC	8.5 kb
295	02603-L02074	Exon 13	4473-4474	AATGGCTGAACT-AGAAGCTGTGTT	0.3 kb
463	11283-L12001	Exon 13	159 nt after exon 13	CTCACAACTAAT-ATACCAGTCAGA	5.7 kb
305	00833-L00349	Exon 14	4648-4649	CCAGAAGGCCTT-TCTGCTGACAAG	2.1 kb
328	00778-L00347	Exon 15	4782-4783	CTCTGGGAGTCT-TCAGAATAGAAA	3.2 kb
337	00779-L00003	Exon 16	4938-4939	ATCTGGAATCAG-CCTCTTCTCTGA	3.5 kb
346	00780-L00283	Exon 17	5246-5247	TTGCCAGAAAAC-ACCACATCACTT	3.7 kb
355	00781-L00284	Exon 18	5326-5327	TTTGTGTGTGAA-CGGACACTGAAA	0.6 kb

364	00782-L00285	Exon 19	5401-5402	ACCCAGTCTATT-AAAGAAAGAAAA	6.3 kb
389	00783-L00356	Exon 20	5462-5463	TGGTCAATGGAA-GAAACCACCAAG	6.0 kb
399	00784-L12004	Exon 21	5536-5537	GAAATCTGTTGC-TATGGGCCCTTC	1.9 kb
407	00785-L00288	Exon 22	5610-5611	TTCTGTGGTGAA-GGAGCTTTCATC	1.5 kb
415	00786-L00289	Exon 23	5654-5655	TCCACCCAATTG-TGGTTGTGCAGC	2.5 kb
427	02831-L13862	Exon 24	6274-6475	TCAATGGAAGGA-GAGTGCTTGGGA	
		Stop codon	5822-5824		

Table 2.5 The *BRCA1* probes included in the SALSA® MLPA® P002-D1 probemix. Indicated are the length of each of the probes, the SALSA MLPA probe number, exon number, the ligation site with the partial ligation surrounding sequences and the distance from the next probe (as described in COA Version 36, issued on 25/02/2015).

Length (nt)	SALSA MLPA probe	*BRCA1* exon	Ligation site	Partial ligation sequence	Distance to next probe
324	18142-L23024	Upstream	4.6 kb upstream exon 1	TCAGGGTCCTTA-AAATAACAGTCT	3.9 kb
166	02808-L25084	Upstream	0.7 kb upstream exon 1, reverse	TCTGCGCACTCG-TAGTTCCACCCC	0.9
154	00763-L22990	1a	192 – 191, reverse	AGCAGAGGGTGA-AGGCCTCCTGAG	0.2 kb
289	20028-L27338	1b	Exon 1b, 207 nt downstream exon 1a	GGGGGCACTGA-GTGTCCGTGGGG	1.0 kb
		Start codon	233-235 (exon 2)		
178	00765-L22993	2	248 – 249	ATTTATCTGCTC-TTCGCGTTGAAG	8.3 kb
421	20035-L22994	3	335 – 336	TCAAGGAACCTG-TCTCCACAAAGT	9.3 kb
190	00767-L22995	5 (4)	393 – 394	ACTTCTCAACCA-GAAGAAAGGGCC	1.6 kb
374	20032-L27342	6 (5)	473 – 474	CGAGATTTAGTC-AACTTGTTGAAG	0.8 kb
220	00769-L22997	7 (6)	637 – 638	AACCGTGCCAAA-AGACTTCTACAG	4.3 kb
403	20034-L27629	8 (7)	718 – 719	CTTGGAACTGTG-AGAACTCTGAGG	2.6 kb
238	01005-L23000	9 (10)	813 – 814	CGTTAATAAGGC-AACTTATTGCAG	1.3 kb
251	00772-L23001	10 (11)	853 – 854	TTGTTACAAATC-ACCCCTCAAGGA	1.1 kb
263	18039-L00345	11 (10)	994 – 995	AAGCGTGCAGCT-GAGAGGCATCCA	0.5 kb
382	20033-L22619	11 (10)	1448 – 1449	AGTCTGAATCAA-ATGCCAAAGTAG	0.4 kb
296	18135-L27339	11 (10)	1838 – 1837, reverse	CGTTTGGTTAGT-TCCCTGATTTAT	0.4 kb
233	18136-L23325	11 (10)	2238 – 2239	CCTACAACTCAT-GGAAGGTAAAGA	0.5 kb
340	20030-L27341	11 (10)	2748 – 2749	TGAAGTTAACCA-CAGTCGGGAAAC	0.5 kb
427	20036-L27344	11 (10)	3265 – 3266	ATGTCACCTGAA-AGAGAAATGGGA	0.5 kb
281	00774-L23003	11 (10)	3810 – 3811	TCCTAGCCCTTT-CACCCATACACA	0.4 kb
142	18139-L22623	11 (10)	4229 – 4230	TCCTAGCCCTTT-CACCCATACACA	0.6 kb
358	20031-L23004	12 (11)	4377 – 4378	CTCTGAAGACTG-CTCAGGGCTATC	8.5 kb
301	02603-L27340	13 (12)	4473 – 474	AATGGCTGAACT-AGAAGCTGTGTT	0.1 kb

202	18290-L23057	13 (12)	4545 – 4546	TGACTCTTCTGC-CCTTGAGGACCT	0.2 kb
459	18169-L23037	13 (12)	159 nt downstream exon 13	CTCACAACTAAT-ATACCAGTCAGA	5.7 kb
269	20027-L27336	14 (13)	4648 – 4649	CCAGAAGGCCTT-TCTGCTGACAAG	2.1 kb
332	00778-L23026	15 (14)	4782 – 4783	CTCTGGGAGTCT-TCAGAATAGAAA	3.2 kb
160	20022-L27333	16 (15)	4938 – 4939	CTCTGGGAGTCT-TCAGAATAGAAA	0.3 kb
196	18144-L22627	16 (15)	5215 – 5216	ACCCCAGAAGAA-TTTGTGAGTGTA	3.3 kb
347	18031-L23028	17 (16)	5246 – 5247	TTGCCAGAAAAC-ACCACATCACTT	3.7 kb
256	20026-L27335	18 (19)	5326 – 5327	TTTGTGTGTGAA-CGGACACTGAAA	0.6 kb
214	20024-L23321	19 (18)	5401 – 5402	ACCCAGTCTATT-AAAGAAAGAAAA	6.3 kb
393	00783-L23319	20 (19)	5462 – 5463	TGGTCAATGGAA-GAAACCACCAAG	6.0 kb
226	20025-L27334	21 (20)	5536 – 5537	GAAATCTGTTGC-TATGGGCCCTTC	1.9 kb
412	00785-L23318	22 (21)	5610 – 5611	TTCTGTGGTGAA-GGAGCTTTCATC	1.5 kb
184	20023-L23035	23 (22)	5654 – 5655	TCCACCCAATTG-TGGTTGTGCAGC	1.9 kb
439	18140-L04795	24 (23)	5722 – 5723	ATGTGTGAGGCA-CCTGTGGTGACC	0.1 kb
		Stop codon	5822-5824 (exon 24)		
149	20021-L27332	24 (23)	5836 – 5837	CTGCAGCCAGCC-ACAGGTACAGAG	0.3 kb
310	20029-L23320	24 (23)	6175 – 6176	GCTGGAAGCACA-GAGTGGCTTGGC	

Table 2.6 The *BRCA1* probes included in the SALSA® MLPA® P087-C1 probemix. Indicated are the length of each of the probes, the SALSA MLPA probe number, exon number, the ligation site with the partial ligation surrounding sequences and the distance from the next probe (as described in COA Version 23, issued on 19/06/2014).

Length (nt)	SALSA MLPA probe	*BRCA1* exon	Ligation site	Partial ligation sequence	Distance to next probe
		Sart codon	233-235 (ex 2)		
160	02808-L25904	Upstream	703 nt before ex 1a, reverse	TCTGCGCACTCG-TAGTTCCACCCC	0.7 kb
148	02807-L01268	Exon 1a	67 nt before ex 1a	TCATCCGGGGGC-AGACTGGGTGGC	0.1 kb
436	02100-L02537	Exon 1a	67-68	CGTGGCAACGGA-AAAGCGCGGGAA	1.4 kb
167	02810-L02239	Exon 2	299-298, rev	GATGGGACACTC-TAAGATTTTCTG	8.3 kb
175	02811-L02240	Exon 3	361-360, rev	ACTTACTTGCAA-AATATGTGGTCA	9.3 kb
346	03395-L12877	Exon 5	3 nt after ex 5, rev	ACCAAATTATAT-ACCTTTTGGTTA	1.6 kb
208	02813-L02242	Exon 6	516-517	TGCTTTTCAGCT-TGACACAGGTTT	0.7 kb
219	02814-L02243	Exon 7	611-610, rev	GTAGCCCATACT-TTGGATGATAGA	4.3 kb
226	02815-L02244	Exon 8	9 nt before exon 8	GTTCTTTACCAT-ACTGTTTAGCAG	2.6 kb
234	02816-L02245	Exon 9	10 nt after ex 9, rev	CAAAGGTTCTCT-TTGACTCACCTG	1.4 kb
355	03822-L03285	Exon 10	884-883, rev	TGCAGAATCCAA-ACTGATTTCATC	2.4 kb
276	02818-L02247	Exon 11	2238-2239	CCTACAACTCAT-GGAAGGTAAAGA	2.5 kb
287	02819-L02248	Exon 12	4400-4399, rev	GGTTAAAATGTC-ACTCTGAGAGGA	8.4 kb
250	03411-L25905	Exon 13	4542-4474	AATGGCTGAACT-AGAAGCTGTGTT	0.1 kb
295	03890-L03337	Exon 13	4542-4243	AAGTGACTCTTC-TGCCCTTGAGGA	5.9 kb
363	11802-L12190	Exon 14	4654-4655	GGCCTTTCTGCT-GACAAGTTTGAG	2.2 kb
329	02821-L02250	Exon 15	4852-4853	CAACAGCTGGAA-GAGTCTGGGCCA	3.2 kb
337	02822-L02251	Exon 16	4996-4997	CCAGAGTCAGCT-CGTGTTGGCAAC	3.5 kb
408	03397-L13116	Exon 17	9 nt after ex 17, rev	TGTAAAGGTTCT-TGGTATACCTGT	3.7 kb
185	03398-L02254	Exon 18	5371-5372	GGAAAATGGGTA-GTTAGCTATTTC	0.5 kb

364	02826-L02255	Exon 19	44 nt before ex 19	TGAAAAGAGCAC-GTTCTTCTGCTG	6.3 kb
390	02827-L02256	Exon 20	27 nt before ex 20	TTTCTCTTATCC-TGATGGGTTGTG	6.1 kb
397	02828-L02257	Exon 21	49 nt after ex 21	CCTTTGTCTTAC-ATAGTGGAGTAT	1.8 kb
200	11457-L12189	Exon 22	5575-5574, rev	TGTACCATCCAT-TCCAGTTGATCT	1.6 kb
416	02830-L02259	Exon 23	21 nt after 23	CTGCATGTACCT-GTGCTATATGGG	1.8 kb
425	04578-L04795	Exon 24	5722-5723	ATGTGTGAGGCA-CCTGTGGTGACC	
		Stop codon	5822-5824 (ex 24)		

Table 2.7 The *BRCA2* probes included in the SALSA® MLPA® P045-B3 probemix. Indicated are the length of each of the probes, the SALSA MLPA probe number, exon number, the ligation site with the partial ligation surrounding sequences and the distance from the next probe (as described in COA Version 33, issued on 09/09/2015).

Length (nt)	SALSA MLPA probe	*BRCA2* exon	Ligation site	Partial ligation sequence	Distance to next probe
160	02143-L09586	FRY gene		GGCCCAGAGTTA-CCGAGTCCTCAC	20.1 kb
		Start codon	228 – 230		
137	02283-L12281	Exon 1	0 – 1	CAGCGCGGGCTT-GTGGCGCGAGCT	0.2 kb
148	02285-L01776	Exon 1	22 nt after exon 1	TGGTAGTGGGTT-GGGACGAGCGCG	0.8 kb
166	02486-L01985	Exon 2	271 – 270, reverse	AGCGTGTCTTAA-AAATTTCAAAAA	2.8 kb
178	01599-L10642	Exon 3	472 – 473	AGCGTGTCTTAA-AAATTTCAAAAA	0.2 kb
172	08898-L09587	Exon 3	107 nt after exon 3	CTGGGCAAATCA-GTCTCTCTGGCC	5.7 kb
197	01600-L04671	Exon 4	569 – 570	AATAGTAGACAT-AAAAGTCTTCGC	1.0 kb
309	09809-L10257	Exon 5	688 – 689	TGTAACACCACA-AAGAGATAAGTC	0.1 kb
346	04585-L03983	Exon 6	728 – 727, reverse	ACAAACTTTGGT-GTATGAAACAAA	0.3 kb
202	08265-L08128	Exon 7	812 – 813	ATGTCTTGGTCA-AGTTCTTTAGCT	2.6 kb
220	01602-L01184	Exon 8	264 nt after exon 8	TCTGACTTTCCA-ACTCATTGTGGA	1.8 kb
229	01603-L01185	Exon 9	1001 – 1002	AACACAAATCAA-AGAGAAGCTGCA	1.6 kb
247	01604-L01186	Exon 10	1374 – 1375	GAAGTGACAAAA-TCTCCAAGGAAG	3.7 kb
256	02279-L01770	Exon 11	2192 – 2193	TCTGAAGAACCA-ACTTTGTCCTTA	4.8 kb
274	01606-L01188	Exon 11	6992 – 6993	TCTCTTTTTACA-TGTCCCGAAAAT	3.3 kb
283	01607-L01189	Exon 12	183 nt before exon 12	AAACAGAACAAA-AATGTAATTGAC	2.5 kb
301	02280-L01771	Exon 13	7216 – 7215	GTACACAGGTAA-TCGGCTCTAAAG	8.2 kb
154	09297-L08066	Exon 14	7394 – 7395	TCTGCTACAAGA-AATGAAAAAATG	1.5 kb
326	01610-L01192	Exon 15	7762 – 7763	CAGTCTGTATCT-TGCAAAAACATC	1.3 kb
337	01611-L01193	Exon 16	7975 – 7976	ACAGTTGGCTGA-TGGTGGATGGCT	4.8 kb
355	02281-L01772	Exon 17	8158 – 8157, reverse	TTAGGCATCTAT-TAGCAAATTCCT	0.8 kb

364	01613-L01195	Exon 18	8482 – 8483	TCAGAAGATTAT-TCTTCATGGAGC	7.0 kb
382	01614-L01196	Exon 19	8602 – 8603	GTATACCAAACT-TGGATTCTTTCC	0.5 kb
391	08266-L08129	Exon 20	8743 – 8744	ATCTGGATTATA-CATATTTCGCAA	5.7 kb
409	02069-L01970	Exon 21	8909 – 8910	ACAAGACAGCAA-GTTCGTGCTTTG	2.7 kb
418	01617-L01199	Exon 22	9100 – 9101	TGCTGAACAAAA-GGAACAAGGTTT	0.3 kb
191	09812-L10643	Exon 23	9214 – 9215	ATCATCAGATTT-ATATTCTCTGTT	0.3 kb
436	08267-L08130	Exon 24	9455 – 9454, reverse	GAAACGACAAAT-CCTATTAGGTCC	14.8 kb
445	08268-L08131	Exon 25	9706 – 9707	AGAGACATTCAA-CAAAATGAAAAA	2.0 kb
463	11984-L15346	Exon 26	9786 – 9787	TACTGCATGCAA-ATGATCCCAAGT	1.3 kb
476	09293-L15678	Exon 27	9988 – 9989	AAAGTCTTGTAA-AGGGGAGAAAGA	0.6 kb
319	09296-L11090	Exon 27	10638 – 10639	ATCGGGCAAAAA-TCGTTTTGCCCG	8.4 kb
		Stop codon	10482 – 10484 (ex 27)		
454	02144-L01619	N4BP2L1 (CG018) gene		CATTATTATTGA-TAATACCAACCT	
265	06800-L02040	HSCB gene/ CHEK2 promoter	2.1 kb before exon 1	TGGCTGAAGAAA-TCTGGGTGGACA	44.0 kb
400	02579-L12282	Exon 9 (10)	1008 – 1009	CTGTTTGACAAA-GTGGTGGGGAAT	4.0 kb
495	01772-L15680	Exon 11 (12)	1173 – 1171, reverse	TGCCCAAAATCA-TAATCTAAAATT	

Table 2.8 The *BRCA2* probes included in the SALSA® MLPA® P077-A3 probemix. Indicated are the length of each of the probes, the SALSA MLPA probe number, exon number, the ligation site with the partial ligation surrounding sequences and the distance from the next probe (as described in COA Version 06, issued on 17/11/2014).

Length (nt)	SALSA MLPA probe	*BRCA2* exon	Ligation site	Partial ligation sequence	Distance to next probe
		Start codon	228-230 (ex 2)		
292	08343-L08275	Upstream	1391 nt before exon 1	ATACTGACCAAT-TTACAGGATTAC	1.5 kb
148	01596-L20889	Exon 1	130-131	GACAGATTTGTG-ACCGGCGCGGTT	0.9 kb
427	12323-L13316	Exon 2	234-235	TAAAAATGCCTA-TTGGATCCAAAG	2.9 kb
256	12305-L14344	Exon 3	64 nt after exon 3	CAAACCTGTGTT-AAAATCTTAGCT	4.4 kb
244	12304-L13297	Exon 3	1261 nt before exon 4	TAGTTACCATCA-ACTATTGGAACC	1.3 kb
178	12294-L13287	Exon 4	626-627	GTTTCCTGTCCA-CTTCTAAATTCT	0.7 kb
337	12313-L13306	Exon 5	206 nt before exon 5	GCTGAAATTTGT-GAGTACATATGT	0.5 kb
154	12290-L13283	Exon 6	79 nt after exon 6	TGCCCTGAGATT-TACAAATCTGTA	0.3 kb
265	12306-L13299	Exon 7	61 nt after exon 7	ATTGTCATCTCT-AATACTTCTGTT	1.9 kb
202	12297-L13290	Intron 7	898 nt before exon 8	CATTGTTTCTCA-AATTGCAGATTA	0.9 kb
391	12319-L13312	Exon 8	888-889	CTGAAACTGTAT-TTCCTCATGATA	0.2 kb
472	12326-L13319	Exon 8	143 nt after exon 8	CTGGTCCTATAT-GTGATTTTAACT	1.3 kb
228	12301-L14436	Exon 9	934-935	TTTTTCCAATCA-TGATGAAAGTCT	2.0 kb
355	12315-L13308	Exon 10	1672-1673	AGACTGCATTCT-TGCAGTAAAGCA	4.4 kb
196	12296-L13289	Exon 11	3225-3226	TCTTAGGTCCAA-TTTCAAATCACA	2.5 kb
136	12289-L20891	Exon 11	5763-5764	CTGCATTTAGGA-TAGCCAGTGGTA	1.3 kb
160	19614-L26252	Exon 11	7066-7067	CCTTATCTTAGT-GGGTAAGTGTTC	3.5 kb
400	12320-L13313	Exon 12	7154-7155	GCTTCAAAAAGC-ACTCCAGATGGT	0.8 kb
214	12299-L13292	Intron 12	770 nt after exon 12	CCCTAAAATGGT-TGTAGATACCTA	1.3 kb
328	12312-L13305	Exon 13	79 nt before exon 13	ATAATTGTCTCA-AATTTTTTGTGT	0.9 kb
346	12314-L13307	Intron 13	788 nt after exon 13	GACCTAGCTCCT-TCCCACACTTGG	7.4 kb
454	12324-L13317	Exon 14	7500-7501	AGTGTGTTAGGA-ATATTAACTTGG	1.4 kb
418	12322-L13315	Exon 15	7729-7730	GAAACAAAGGCA-ACGCGTCTTTCC	1.3 kb

364	12316-L13309	Exon 16	7915-7916	TTTTCAGTTTCA-CACTGAAGATTA	0.3 kb
221	12300-L13293	Exon 16	159 nt after exon 16	TGTAGAAGTCTT-TTGAAAAGTGCT	4.7 kb
238	12303-L13296	Exon 17	82 nt after exon 17	TTCCATGTCAAA-ATGTTGCACAAG	0.4 kb
463	12325-L13318	Exon 18	26 nt before exon 18	ACACTTCCTAAA-ATATGCATTTTT	0.7 kb
274	12307-L13300	Exon 18	275 nt after exon 18	TTGCTGCCCTCT-TGTTCTCATAGC	6.7 kb
232	12302-L13295	Exon 19	8651-8652	TTATCATCGCTT-TTCAGTGATGGA	0.6 kb
409	12321-L13314	Exon 20	5 nt after exon 20	CATGAAGGTAAA-ATTAGTTATATG	5.7 kb
301	12309-L13302	Exon 21	31 nt after exon 21	TATAATGAGGCT-TGATGATTATTC	1.2 kb
382	12318-L13311	Intron 21	1180 nt after exon 21, 1345 nt before exon 22	CTGTTCCGAAAT-GTTACAATTTGG	1.8 kb
191	09811-L10259	Exon 23	4 nt before exon 23	ATCTTTCTCCAA-ACAGTTATACTG	0.3 kb
436	01618-L14536	Exon 24	9407-9408	TTCAGCAAATTT-TTAGATCCAGAC	14.8 kb
283	12308-L13301	Exon 25	9677-9678	TCTGCTAGTCCA-AAAGAGGGCCAC	2.1 kb
184	12295-L13288	Exon 26	9841-9842	GCCGTACACTGC-TCAAATCATTCC	1.4 kb
166	12292-L14535	Exon 27	10111-10112	TCCGGCTGCACA-GAAGGCATTTCA	0.2 kb
310	12310-L13303	Exon 27	10289-10290	GCTCTTTTGTCT-GGTTCAACAGGA	
		Stop codon	10482-10484 (ex 27)		

2.3.4.1 Hybridization and ligation of the probes

Five microliters of genomic DNA or $T._1E$ buffer in the case of the no DNA control, was denatured for 5 min at 98°C using the GeneAmp® system 9600 (Applied Biosystems). After denaturation, the samples were rapidly cooled to 25°C before adding the hybridization master mix (1.5 µl SALSA MLPA buffer® [KCl, Tris-HCl, EDTA, PEG-6000, oligonucleotide as per package insert] per reaction mixed beforehand with 1.5 µl SALSA® MLPA® probemix per reaction) and carefully mixed by pipetting. The mixture was denatured for 1 min at 95°C, where after the reactions were incubated at 60°C for 16 hours.

Following overnight incubation, the temperature was cooled to 54°C and the instrument paused. The ligase master mix consisting of 25 µl dH_2O mixed with 3 µl Ligase Buffer A® (10 X Coenzyme NAD of bacterial origin), 3 µl Ligase Buffer B® (Tris-HCl, $MgCl_2$, non-ionic detergent) and 1 µl Ligase-65 enzyme® (Glycerol, EDTA, Beta-Mercaptoethanol, KCl, Tris-HCl, nonionic detergent Ligase-65 enzyme of bacterial origin) (prepared beforehand for each reaction) was added to each tube. The reactions were ligated at 54°C for 15 min, followed by the inactivation of the enzyme at 98°C for 5 min. After denaturation, the samples were rapidly cooled to 20°C and paused.

2.3.4.2 PCR reaction

Ten microliters of well mixed master mix comprising of 7.5 µl dH_2O, 2 µl SALSA PCR primer mix® (synthetic oligonucleotides labelled with fluorescent dye [FAM or Cy5], dNTPs, Tris-HCl, KCl, EDTA and non-ionic detergent) and 0.5 µl SALSA Polymerase® (Glycerol, non-ionic detergents, EDTA, DTT, KCl, Tris-HCl, Polymerase enzyme of bacterial origin) were added to the PCR tubes containing 40 µl of ligated products at room temperature. The products were gently mixed by pipetting, briefly centrifuged and returned to the conventional PCR system (GeneAmp® system 9600, Applied Biosystem) for amplification. The amplification regime was as follows: 35 cycles of 95°C for 30 sec, 60°C for 30 sec and 72°C for 60 sec. The PCR program was concluded with a final incubation at 72°C for 20 min, where after the samples were cooled to 15°C and the program paused.

The PCR products were stored in the dark at 4°C until separation by capillary electrophoresis. For longer periods, the products were stored at -25°C to - 15°C.

2.3.4.3 Fragment analysis

A small volume of the amplified products (1 µl) was mixed with 8.5 µl Hi-Di™ formamide (Applied Biosystems, US) together with 0.5 µl of GeneScan-500 LIZ standard (Applied Biosystems). The samples were mixed, briefly centrifuged, denatured for 5 min at 98 °C, snap cooled on ice for 5 min and electrophoresed at 15 kV utilizing POP-7 (ABI-3130) polymer with a filter set G5. Fragment electrophoresis was performed on an Applied Biosystems 3130 XL Genetic Analyzer. The files were generated as peaks for the various exons and included control peaks for the ligated probes.

2.3.4.4 MLPA data analysis

Multiplex ligation-dependent probe amplification data was analyzed using GeneMarker® software v 2.6.4 (SoftGenetics, LLC, State College PA, USA). The program was used to analyze the exon copy number or dosage quotient (DQ) for each ligated product. Once the raw data files were received from the Genetic Analyzer, it was uploaded into the program and visually inspected by generating a synthetic gel image. Depending on the image, various correction icons were selected to enhance and clean the raw data image. Once optimal, the raw data files were ready to be processed (GeneMarker®, The Biologist Friendly Software User Manual 2014, SoftGenetics®).

Analysis included the application of a sizing standard, filtering of noisy peaks and comparison of the data to a known allelic panel. The application of the sizing standard is essential due to differential fragment mobility during capillary gel electrophoresis. This process was followed according to the manufacturer's manual (GeneMarker®, The Biologist Friendly Software User Manual 2014). Once the data was sized, the filtering parameters applied and a reference panel selected, the data could be analyzed.

Upon analysis, raw data is divided into two parts, one is a synthetic gel image which displayed the unprocessed data in a traditional gel format with smaller fragments on the left and larger fragments on the right. The second one is an electropherogram which displays fluorescent signal intensities as a single line trace for each dye color. Each generated MLPA fragment differs in size standard and for BRCA SALSA® probes used to screen patients, a size standard GS500 was used for fragment analysis. Accordingly, MRC-Holland (www.mrc-holland.com) has designed MLPA panels covering the entire screened fragments and the panel used for BRCA SALSA® probes were unique for the specific gene. To determine the presence of deletions and duplications using MLPA, the data were visualized and analyzed using four steps. These included a sample list showing samples in the dataset and their respective MLPA score, a trace overlay which displayed normalized traces of samples where the reference probe is appearing in light red and the sample appearing as a blue peak. Furthermore, it contains a dosage histogram where the bar appear larger in case of a detected duplication and smaller or absent for deletions, ratio plots displaying in a graphical form the ratio of normalized peak intensities between the reference and the sample and finally the report table displaying the probe or exon name. The resulting data were analyzed using the GeneMarker® software. Visual peak pattern evaluation was carried out following the manufacturer's recommendations.

Chapter 3

Results and Discussion

Copy number variations could play a major role in the inheritance of BC within the SA population, as many BC patients with a strong family history of the disease are only carriers of benign polymorphisms. As MLPA is based on the amplification of up to 60 probes within a single tube (each detecting a specific DNA sequence), optimization and validation of the technique was critical. Once validated, the patients could be analysed regarding the presence of possible deletions or duplications present within the familial BC genes.

3.1 MLPA Optimization

3.1.1 MLPA quality control fragments

Various quality parameters or scores have been built into the multi-step technique (in the form of control fragments) and the analysis software to indicate the degree of accuracy of potential results. The quality of the control fragments reflected the quality of the run and therefore the patient data. These control fragments formed part of the intra-sample normalization and quality check.

A total of nine control fragments are included in each of the SALSA® MLPA® probe sets used (MLPA® DNA Protocol version MDP-005, revised 22 September 2014). These consist of a benchmark probe of 92 nt, four Q-fragments (64, 70, 76 and 82 nt), two D-fragments (88 and 96 nt) and two gender specific probes (X and Y) (Fig. 3.1 A & B). Each of these control fragments reflected the quality of a certain section of the MLPA procedure.

The benchmark probe (at 92 nt) is a normal probe, which should always be present. This probe is the one to which all the other quality control probes are measured

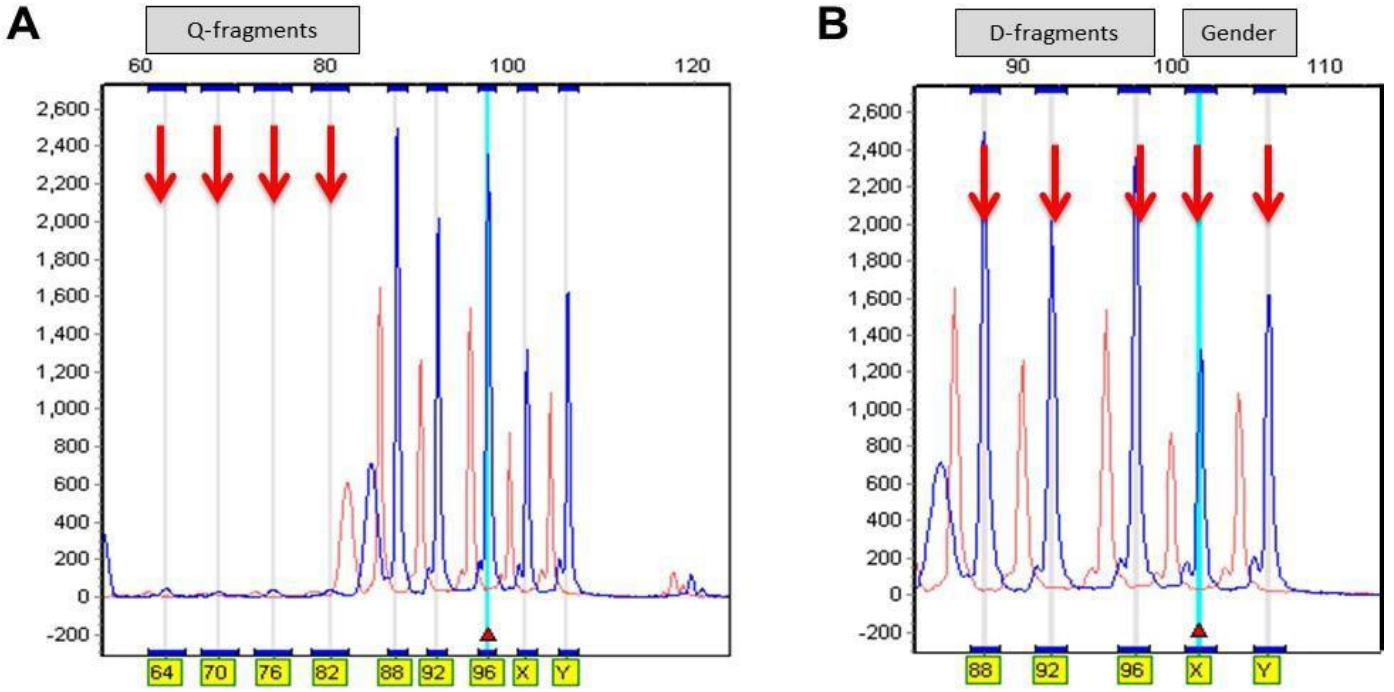

Figure 3.1 MLPA analysis, highlighting the presence of the nine quality control fragments present for this male individual. **A.** The four Q-fragments positioned at 64, 70, 76 and 82 nt respectively (as indicated by the red arrows) were shadowed when the total DNA concentration was above 100 ng. **B.** The two D-fragments (positioned at 88 and 96 nt), the benchmark probe (92 nt) and the two gender specific fragments (X and Y) were included in each run (indicated by the red arrows). The size of each fragment is highlighted in yellow.

against (Fig. 3.1 A & B). The four Q-fragments (64, 70, 76 and 82 nt) indicated whether sufficient amounts of DNA was present within each of the reactions and whether ligation occurred (Fig. 3.1 A). These fragments are complete and do not have to hybridize. Within an optimal reaction, these fragments should ideally be less than 33% of the 92 nt control fragment. The height of these fragments decreases when more sample DNA is added to the reactions as evident from Figure 3.1 A.

The height of the two D-fragments (88 and 96 nt) indicate whether DNA denaturation of the template was optimal. Denaturation can be influenced by the presence of additional salt in the DNA sample or the presence of GC- rich areas within some genes. These two fragments detect sequences in exceptionally strong CpG islands, which is in most cases difficult to denature. It is recommended that the height of these peaks should be more than 40% of the 92 nt control fragment (MLPA® DNA Protocol version MDP-005, revised 22 September 2014; Fig. 3.1 B).

The last of the quality fragments are the X and Y fragments (100 and 105 nt). These fragments indicate the sex of the patient and are valuable as it can indicate potential sample swapping (Fig. 3.1 B). Once a run has been performed, the quality of the run was critically evaluated using all nine quality parameters. If some of these parameters were out of range, the run was repeated and the initial data discarded.

3.1.2 DNA quality

Good quality DNA is essential for accurate MLPA data analysis as impurities such as salts, phenol or ethanol will influence the intra-sample normalization. All the DNA samples analysed in this study (test and reference samples) were extracted using the salting-out technique, as recommended by MRC Holland (MLPA® DNA Protocol version MDP-005, revised 22 September 2014). An additional 70% EtOH wash step was performed to remove extra impurities. The DNA concentrations were meticulously measured using the NanoDrop® ND-100 where after dilution commenced. The DNA dilutions were diluted the day before analysis (as described in the previous chapter) using T.$_1$E (10 mM Tris-HCl [pH 8.2], 0.1 mM EDTA) to

prevent depurination during the initial heat treatments. The dilutions were incubated overnight at 37°C.

3.1.3 DNA quantity

As a variable, starting DNA concentration of 10 – 250 ng in total is indicated, an experiment was performed to determine the optimal amount of DNA to be used in each reaction. Two DNA concentrations were selected representing the same patient and the results were compared, namely 200 ng/µl (a total of 1000 ng as recommended by an accredited Canadian diagnostic laboratory by personal communication) versus 50 ng/µl (total of 250 ng of DNA as indicated by the package insert) (Fig. 3.2).

From the data presented in Fig. 3.2, it is clear that both DNA concentrations delivered data, although the quality differed. For both reactions, the Q-fragments were very low indicating sufficient amounts of DNA (the four fragments are enclosed in the red circle, Fig. 3.2 A & B). As the D-fragments directly to the right of the circle were higher than 40% of the 92 nt control fragment, denaturation of the template proved to be sufficient. The absence of the Y probe confirmed that the individual tested was female.

It is clear that although the 200 ng/µl (total amount of 1000 ng DNA) reaction passed all the quality checks, it produced data with a much lower relative dye signal strength (approximately 1000, Fig. 3.2 A) compared to Fig. 3.2 B with 50 ng/µl (total amount of 250 ng). The relative dye signal strength for the reaction using the lesser amount of DNA was more than twice that indicated in Fig. 3.2 A.

The effect of the use of excess DNA can clearly be seen in Figure 3.3. The excess DNA resulted in less than optimal denaturation of the sample, which is reflected by the lower height of both the D-fragments (88 and 96 nt), compared to that for the 250 ng DNA sample (Fig. 3.3 A). Based on these results, a total amount of 250 ng was used for all test and reference samples within this study.

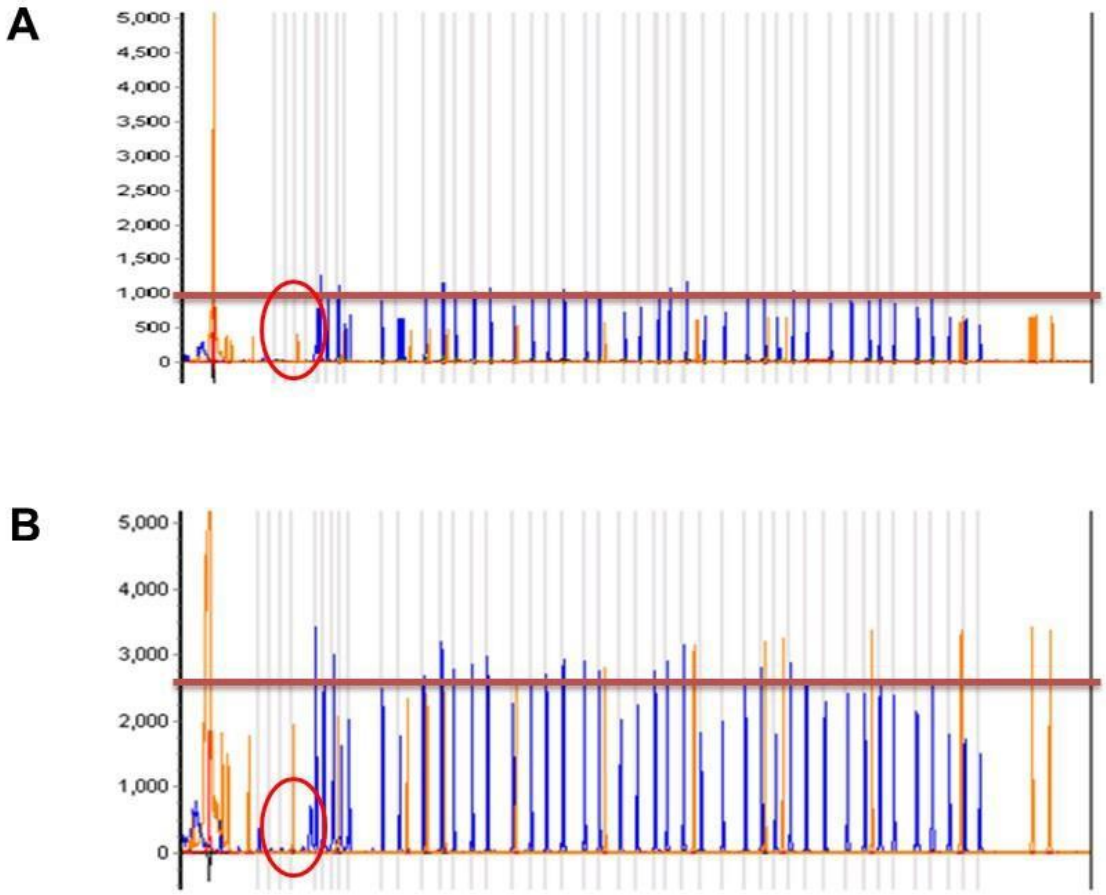

Figure 3.2 The influence of the initial DNA concentration used within an MLPA reaction. For both reactions, the height of the Q-fragments were very low (indicated by red circle) when compared to that of the probes. **A.** A dye signal strength of approximately 1000 is achieved with an initial DNA concentration of 200 ng/µl (total amount of 1000 ng). **B.** A dye signal strength of approximately 2600 is achieved using 50 ng/µl (total amount of DNA is 250 ng).

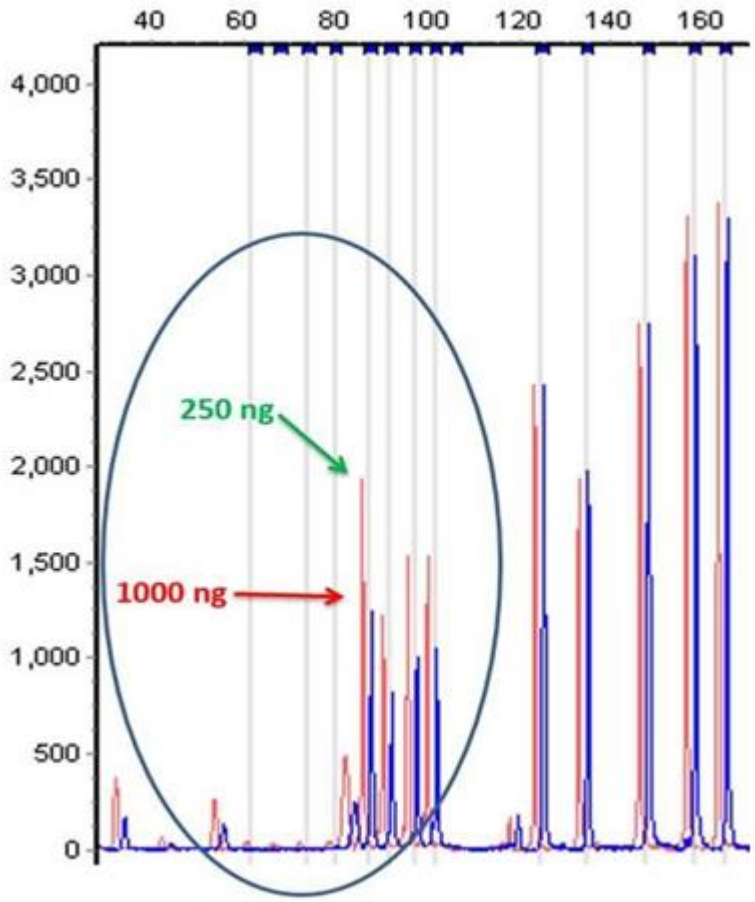

Figure 3.3 The effect of DNA quantity on the denaturation of the sample. Although the reaction worked for both MLPA reactions (Q-fragments are shadowed for both on the baseline) and the probe signals have a good relative dye signal strength, a difference in the height of the D-fragments were observed (highlighted by the blue circle). The height of both the 88 and 96 nt peaks for the sample containing 1000 ng of DNA indicate insufficient denaturation of the template DNA, compared to that of the 250 ng reaction.

3.1.4 Control reactions

Various control reactions were included in every run. These included a No Template Control (NTC), which served to indicate whether any contamination of the master mixes and probes have occurred. The NTC contained no DNA as it was replaced with $T._1E$ (Fig. 3.4 A). All the quality control Q-fragments were present as they are not dependent on hybridization (Fig. 3.4 A). Three DNA samples (one male and two females) representing non-affected individuals without a family history of BC acted as reference controls and were present in each run to assist with the inter normalization between samples.

3.1.5 Sloping of the raw data

Visual inspection of the raw data (un-manipulated data of electropherograms that contains information about all channels of the capillary electrophoresis instrument) was performed after uploading it into the GeneMarker® software v 2.6.4 (SoftGenetics, LLC, State College PA, USA). The data on the graph tended to slope downwards towards the right hand side of the graph (probes increase in nt number), with the shortest probes listed first followed by the longer probes (Fig. 3.5 A). The sloping of the data was due to the shortest probes amplifying more efficiently than the longer probes. This resulted in the shorter probes exhibiting a higher relative dye signal strength (Fig. 3.5 A). The sloping of the data in the MLPA fragments is considered normal when the signal intensity of the largest MLPA fragments is at least approximately 1/3 of the signal intensity of the shortest MLPA fragments and preferably above that for the size standard (Fig. 3.5 A).

Initially, the sloping of the MLPA fragments probes was too steep during the optimization process, with the larger fragments being 1/7 of that of the smaller MLPA fragments (Fig. 3.5 B). The signal intensity of the longer probes was also below that of the sizing standard (Fig. 3.5 B). The steep slope was either due to evaporation of the sample during the overnight hybridization process (lid of the tubes not closing properly) or evaporation during the pipetting of the ligation reaction (occurs at 54°C). All the lids of the tubes were initially opened at once while the tubes were still in the heated block of the PCR instrument, which could have resulted in the evaporation of the already small starting volume (5 µl in total) within each of

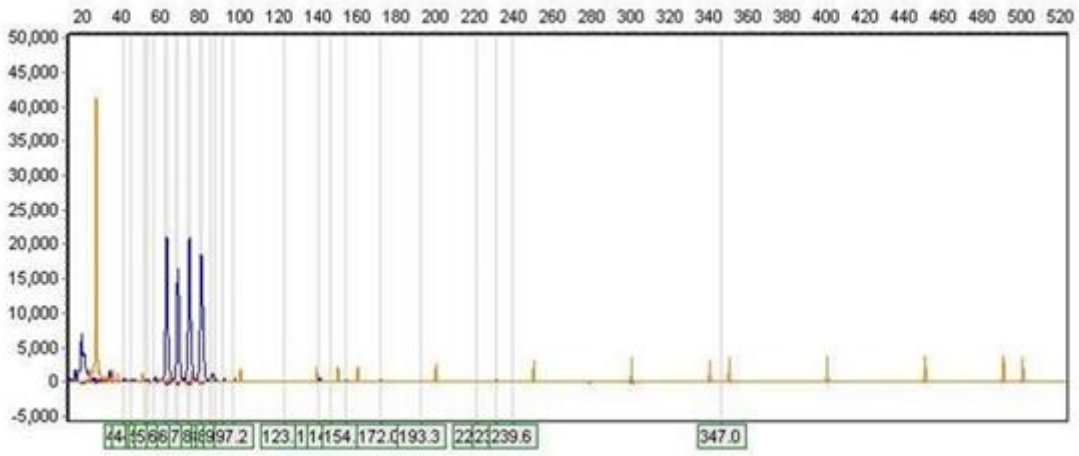

Figure 3.4 A NTC was included in each MLPA run in order to detect potential contamination of MLPA reagents. Within the NTC, the Q-fragments (64, 70, 76 and 82 nt) are clearly visible (dye signal strength of 15 000 – 20 000), with no other peaks visible.

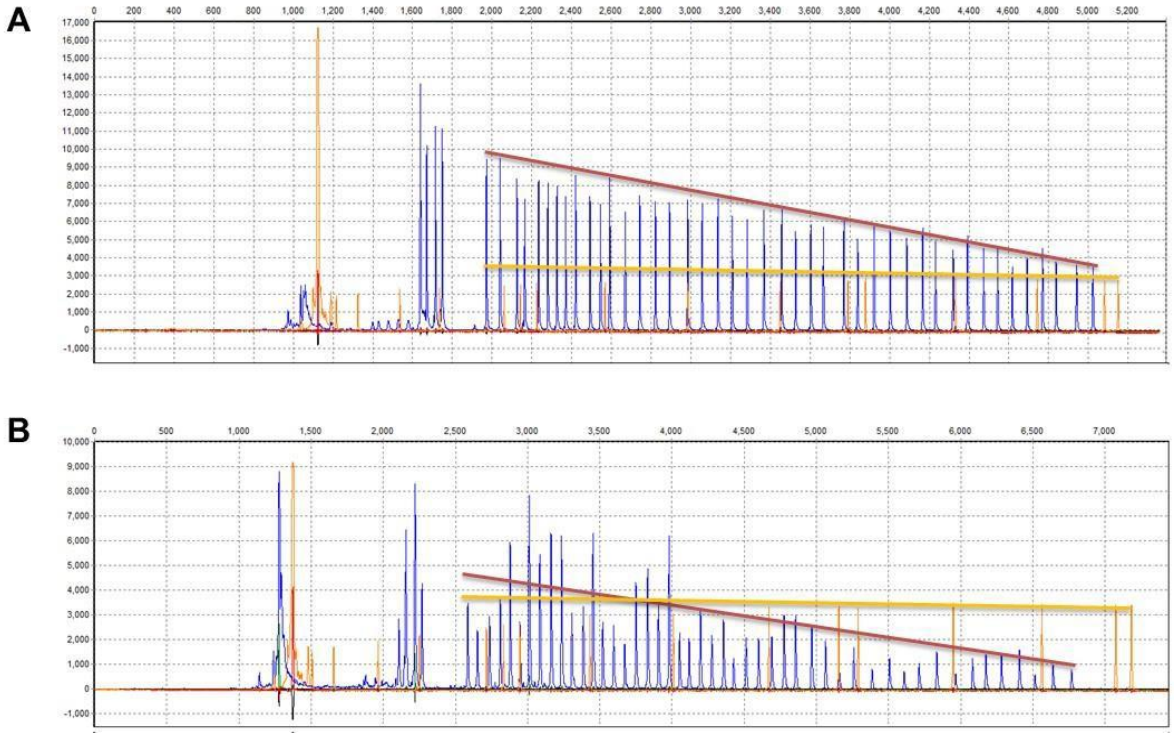

Figure 3.5 The influence of evaporation on the raw data of an MLPA reaction. **A.** Normal sloping of the data present within an optimized reaction, with the preferential amplification of the shorter probes at the start of the slope, gradually going down towards the longer probes on the right hand side. Normal sloping (indicated by the red line) was present when the signal intensity of the longer probes on the right was approximately a third of that of the shorter probes, but still above the fluorescence of the sizing standard (yellow line). **B.** Evaporation of the MLPA reactions occurred when the signal intensity of the longer fragments fell beneath that of the sizing standard. The signal intensity of the longer fragments (towards the right hand side of the graph) was also less than a third of that of the shorter fragments.

the PCR tubes. A preventative strategy was implemented by opening only four tubes at a time during the addition of the MLPA reagents at 54°C. Care was taken to work as quickly and accurately as possible, to prevent possible time delays while performing the technique. The implementation of this step proved to prevent the problems of evaporation and resulted in a normal slope being present when the raw data was inspected (Fig. 3.5 A).

3.1.6 Intra and inter sample normalization

The ultimate aim of performing MLPA was to determine exon copy number or DQ (a method for discriminating normal, deletion, and duplication) for each of the ligation products representing each of the familial BC genes. Before the final analysis, intra-sample normalization was performed during which each probe peak was compared to the peaks of the reference probes within each sample (Fig. 3.6 A). The reference probes detected sequences that were expected to have a normal copy number in all the samples (Fig. 3.6 B). Once the data was normalized for each specific sample, the final probe ratios were determined by comparing the relative probe peaks in the DNA sample to all reference and negative control samples (Fig. 3.6 B). As the reference DNA samples are expected to have a normal copy number, differences in the sample of interest could easily be identified. This action served as inter-sample normalization, where the DQ value represented the copy number. A DQ of 0.7 or below was interpreted as a deletion (as recommended by the GeneMarker® software v 2.6.4), whereas a DQ of 1.3 or above was considered to be a duplication. The MLPA peak pattern of an MLPA negative sample will be identical to that of the reference samples.

3.2 Validation of the MLPA technique

Once all the internal quality parameters of the technique were tested and optimized, validation of the technique commenced. Thanks to the existence of a collaboration agreement between the Division of Human Genetics in Bloemfontein and the Department of Oncology and Human Genetics at McGill University in Canada, 12 SA (CAM2235, 2245, 2262, 2336, 2339, 2362, 2516, 2520, 2645, 2711, 2752 and CAM2975) samples previously analyzed at their facility, together with two Canadian positive controls (CAM2795 and 2951), were included in the validation study. The 12 BC patients all had a positive family history and tested negative for the presence

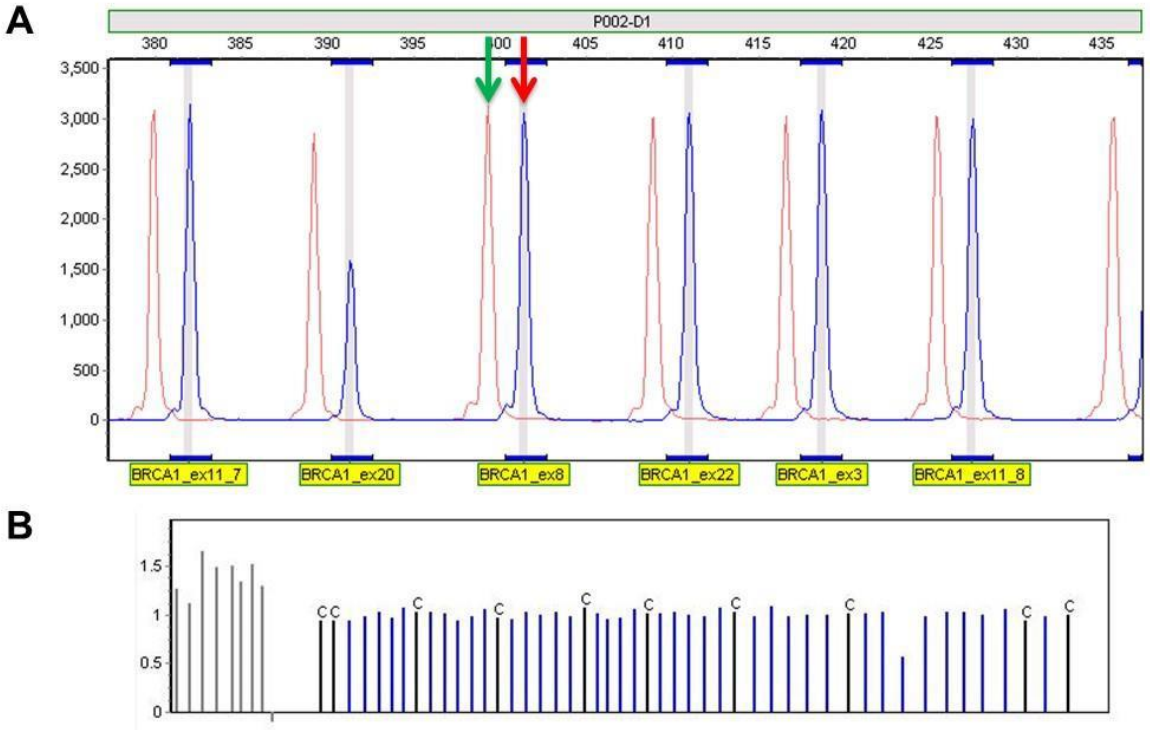

Figure 3.6 Intra-sample and inter-sample normalization of each MLPA reaction was performed to determine the presence of possible deletions or duplications. **A.** Intra-sample normalization of each probe (blue peaks indicated by the red arrow) within the reaction to that of a reference probe set (red peaks indicated by green arrow). **B.** Inter-sample normalization during which the final probe ratios for a specific individual were determined by comparison of each relative probe peak to that of all the reference samples within the run. The reference probe sequences are indicated by the C and are in black.

of smaller deletions/duplications or nonsense mutations within the *BRCA* genes. These patients were screened using HRMA and the protein truncation test. The two Canadian positive controls, tested in the respective accredited laboratories, formed part of this validation study.

3.2.1 MLPA positive Canadian patient (*BRCA1* exon1-2dup)

The DNA of a confirmed MLPA positive Canadian BC patient (CAM2795) carrying the *BRCA1* exon1-2dup mutation (g.90012_97270dup, Del Valle *et al.*, 2010), was collected from McGill University in order to include and assist with the validation study in SA. The 14 samples (two Canadian positives and 12 SA negative samples) included for the validation were initially analysed using the SALSA® MLPA® P002-C1 probe set (Fig. 3.7 A). The intra-sample normalization indicated a duplication of three probes situated within *BRCA1*, for the Canadian patient (Fig. 3.7 A). The probes represented *BRCA1* probe 00763-L00268 sized 148 nt (exon 1A), the *BRCA1* probe 00764-L00269 sized 157 nt (exon 1B) and *BRCA1* probe 00765-L00270 sized 166 nt (exon 2) (Table 2.4). The ligation sites for these probes were located at sequences differing from the pseudo-*BRCA1* exons (Brown *et al.*, 1995). The three duplicated probes remained once the inter-sample normalization were completed (Fig. 3.7 B). The graphical representation (Fig. 3.7 C & D) indicated identical results obtained by both the Canadian (performed during 2014) and the SA laboratory (performed during 2015). The placement of the three probes above the upper horizontal line (DQ of 1.3 or more peak ratio) indicated the presence of a duplication involving consecutive exons 1a, 1b and exon 2 of *BRCA1*. The replication of the Canadian results (Fig. 3.7 C) demonstrated the successful optimization of the SALSA® MLPA® P002-C1 probe set in the SA laboratory (Fig. 3.7 D).

As mutations or polymorphisms in the target sequence can be detected by the probes and cause a reduction in the relative peak height, it is crucial to confirm all potential positive results using a second confirmation probe set or alternative method. The presence of the duplication was confirmed when the sample was screened using the SALSA® MLPA® P002-D1 probe set for *BRCA1*, the follow-up version of SALSA® MLPA® P002-C1 (Fig. 3.8). This probemix has been modified

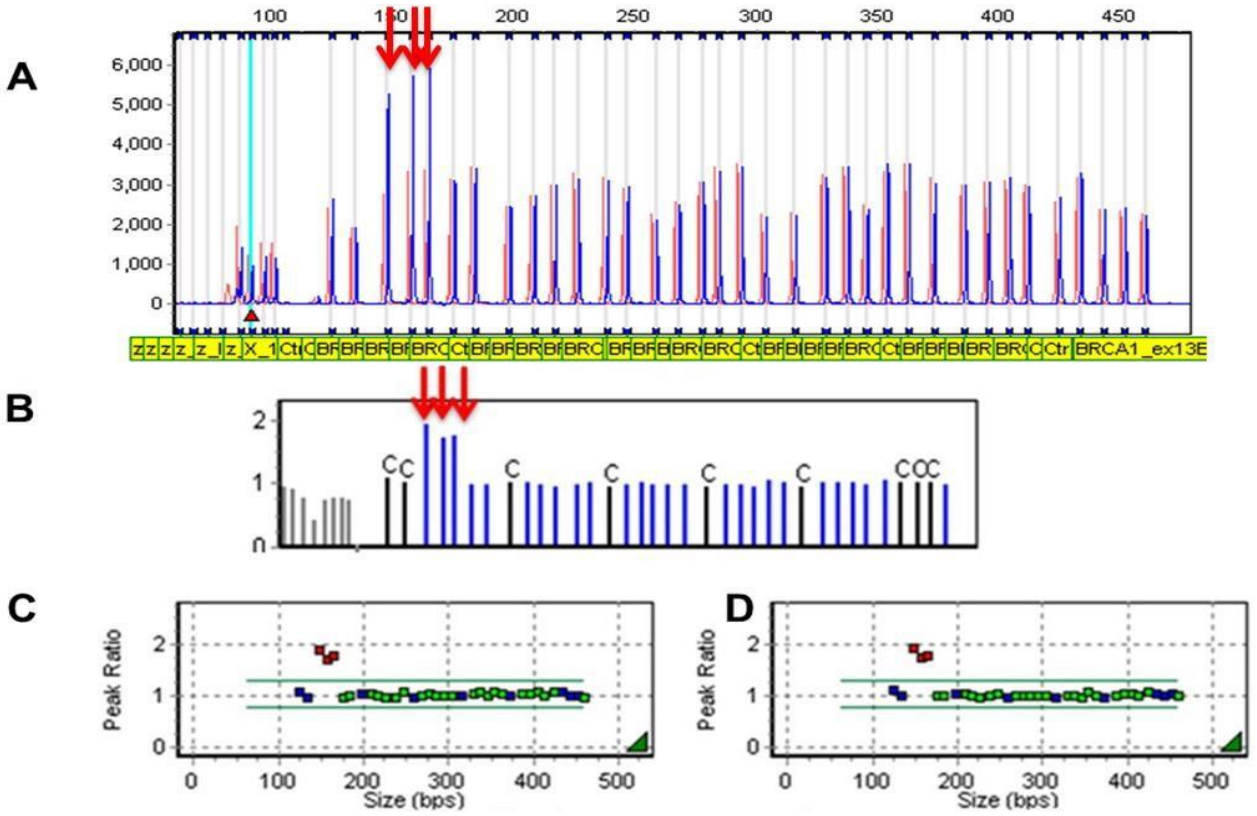

Figure 3.7 Validation results for the Canadian BC patient (CAM2795) carrying the *BRCA1* exon 1-2dup mutation using the SALSA® MLPA® probemix P002-C1. **A.** Raw data indicating a duplication of exons 1a, 1b and 2 (indicated by the red arrows) after intra-sample normalization. **B.** Dosage histogram representing inter-sample normalization indicating the duplication of three probes (indicated by the red arrows). **C**. Graphical representation of the results obtained by the Canadian laboratory (DQ value above 1.3). **D**. Graphical representation of the results obtained for the Canadian patient CAM2795 by the SA laboratory, indicating identical results, namely the duplication of exons 1a, 1b and exon 2 situated within *BRCA1* (DQ value above 1.3).

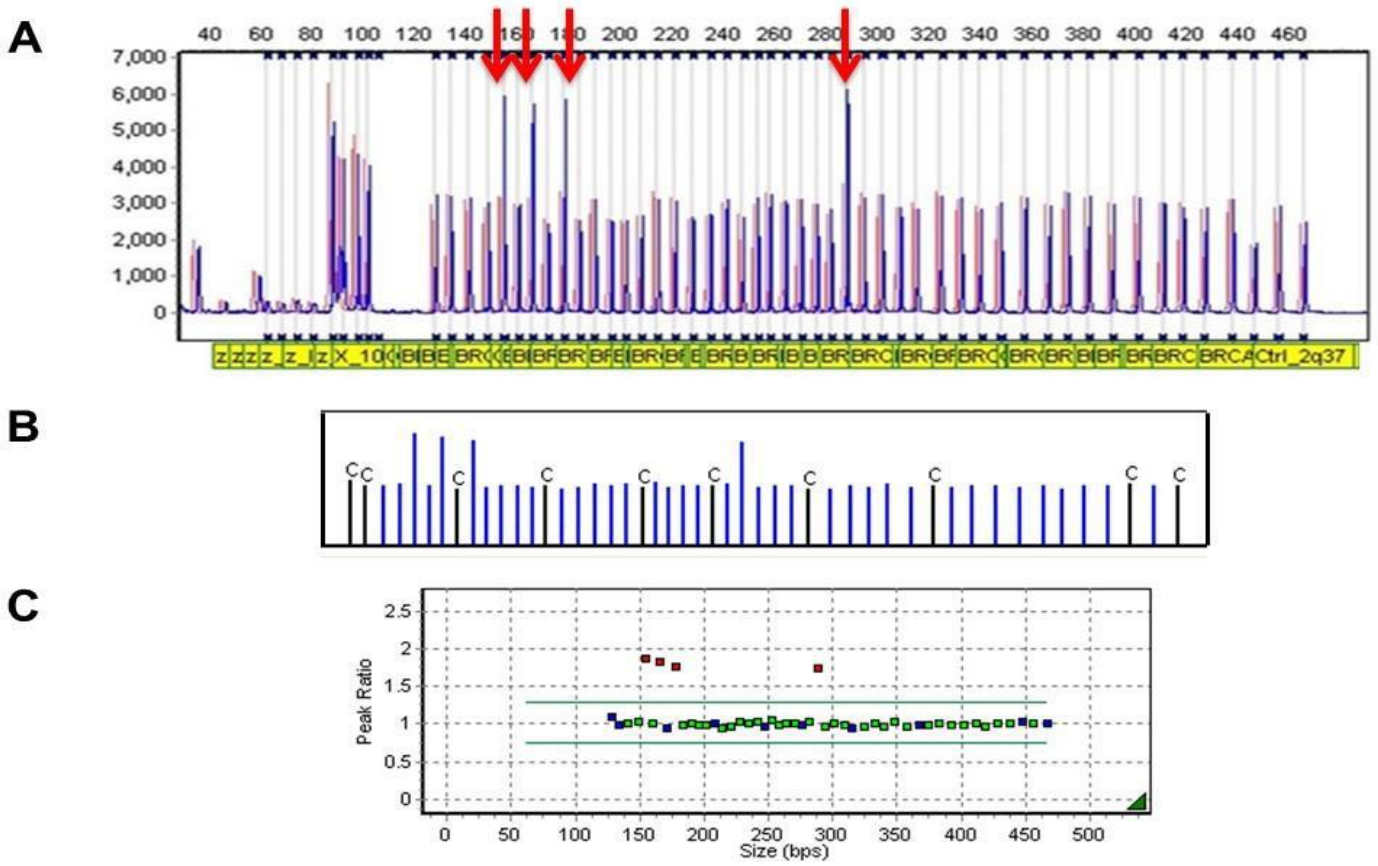

Figure 3.8 Confirmation of the *BRCA1* duplication detected in the Canadian BC patient (CAM2795) using the SALSA® MLPA® P002-D1 probemix. **A.** Raw data indicating a possible duplication of four probes (exons 1a, 1b, 2 and an area upstream of exon 1 indicated by the red arrows) after intra-sample normalization. **B.** Dosage histogram indicating the duplication of four probes for *BRCA1*. **C**. Graphical representation of the results obtained for CAM2795 by the SA laboratory confirming the duplication of *BRCA1* exons 1a, 1b and 2.

as it contained 12 extra *BRCA1* probes and three probes specific for exon 24 (Product Description version 36; issued 25-02-2015) (Table 2.5).

In contrast to previous results, the intra-sample normalization indicated a duplication of four probes situated within *BRCA1* for the Canadian patient (Fig. 3.8 A & B). The probes represented *BRCA1* probe 00763-L22990 sized 154 nt (exon 1A), the *BRCA1* probe 02808-L25084 sized 166 nt (upstream probe), *BRCA1* probe 00765-L22993 sized 178 nt (exon 2) and *BRCA1* probe 20028-L27338 sized 289 nt (exon 1B) (Table 2.5). The four duplicated probes remained once the inter-sample normalization were completed (Fig. 3.8 B & C). The discrepancy in the number of duplicated probes detected between the two probemixes was due to the inclusion of an upstream fourth probe *BRCA1* probe 02808-L25084 (sized 166 nt) (Table 2.5). This probe served as a reference probe and was located 0.7 kb upstream of exon 1 (Product Description version 36; issued 25-02-2015). Utilization of the second SALSA® MLPA® P002-D1 probemix performed by the SA laboratory, therefore confirmed the presence of the mutation within Canadian CAM2795.

3.2.2 MLPA positive Canadian patient (*BRCA1* ins6kbEx13 dup)

The second Canadian MLPA mutation positive patient (CAM2951) was screened within the Canadian laboratory during 2014 using the SALSA® MLPA® P002-C1 probe set (Table 2.4). The sample was retested in the SA laboratory during 2015 using the SALSA® MLPA® P002-D1 probe set for *BRCA1* (Fig. 3.9) (Table 2.5). CAM2951 tested positive for the common and recurrent *BRCA1* ins6kbEx13 duplication (The *BRCA1* Exon 13 Duplication Screening Group, 2010). This duplication involves a 6 kb duplication of exon 13 (Puget *et al.*, 1999). The duplication creates a frameshift in the coding sequence of the gene and has been described for various populations such as Australia, Great Britain and the Netherlands (The *BRCA1* Exon 13 Duplication Screening Group, 2010).

The intra-sample normalization indicated a duplication of three probes situated within *BRCA1* (Fig. 3.9 A). All the probes represented *BRCA1* exon 13, namely probe 18290-L23057 sized 202 nt, probe 02603-L27340 sized 301 nt and probe

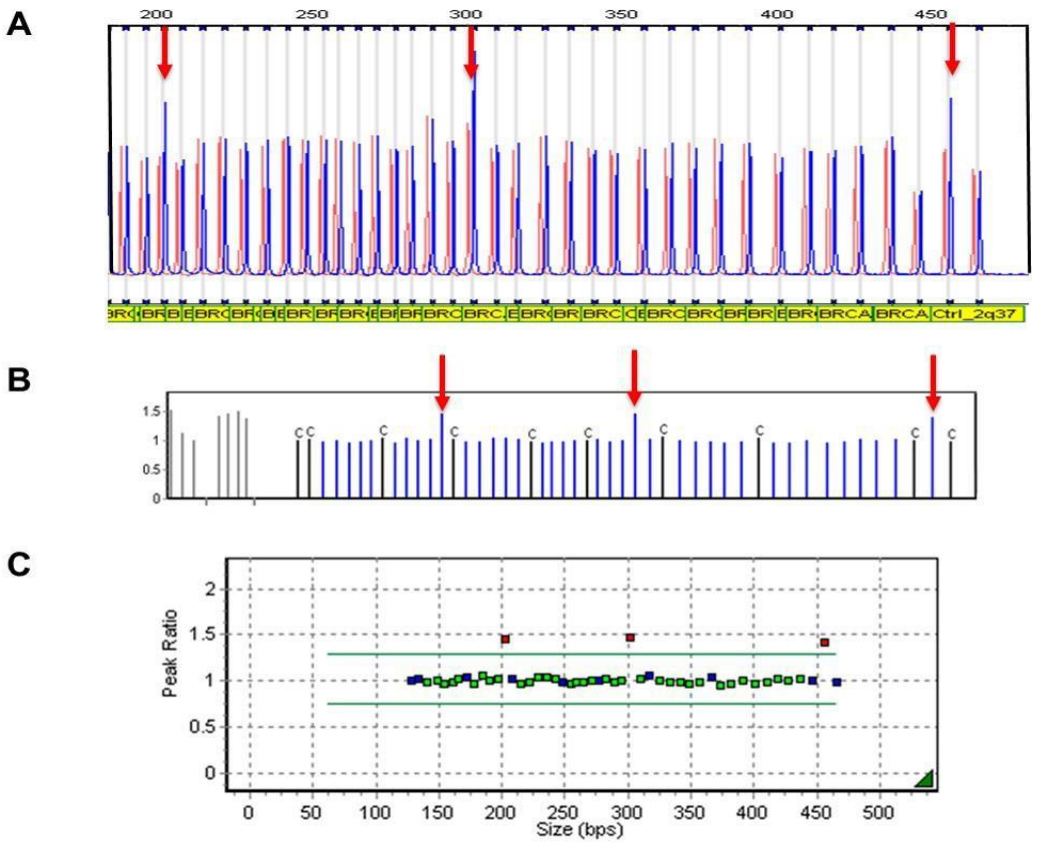

Figure 3.9 Validation results for the second MLPA positive Canadian BC patient (CAM2951) carrying the *BRCA1* ins6kbEx13 duplication. **A.** Raw data for SALSA® MLPA® P002-D1 indicating the duplication of three probes (indicated by the red arrows) all representing exon 13 after intra-sample normalization. **B.** Dosage histogram representing inter-sample normalization indicating the duplication of three probes (indicated by the red arrows). **C**. Graphical representation of the results obtained for CAM2951, with three probes having a DQ value above 1.3.

18169-L23037 sized 459 nt (Table 2.5). The three duplicated probes remained once the inter-sample normalization were completed (Fig. 3.9 B) and was confirmed by the graphical representation (Fig. 3.9 C).

3.2.3 South African MLPA negative BC patients

Twelve SA patients were screened by both parties, once in Canada (during MLPA training session on MLPA) as well as in the SA laboratory here in Bloemfontein. These patients included CAM2235, 2245, 2262, 2336, 2339, 2362, 2516, 2520, 2645, 2711, 2752 and 2975. The results for all the samples corresponded between the two laboratories, with all 12 samples testing negative (DQ values between 0.7 and 1.3) for all the probes sets representing *BRCA1* (Fig. 3.10 A – C) and *BRCA2* (Fig. 3.11 A – C), including both confirmation sets (Tables 2.6 & 2.8). As both positive controls (CAM2795 and CAM2951), together with the results for 12 negative SA BC patients concurred, the validation was deemed successful and the technique could be implemented onto the diagnostic platform of the laboratory.

3.3 Influence of smaller rearrangements on MLPA data

One significant limitation of the MLPA technique relates to the fact that small point mutations within the target sequence of the respective probes could result in false positive results. Mutations or single base changes as far as 20 bp from the ligation site of the probes can result in a reduction in the probe signal, thereby indicating a false duplication or deletion. This reduction could be due to the prevention of ligation of the probe oligonucleotides or due to the destabilizing of the binding of a probe to the sample DNA (MLPA® DNA Protocol version MDP-005; last revised on 22 Sept 2014). As both these genes are highly polymorphic with no specific mutational hotspots, small mutations have been recorded throughout the coding regions of each (BIC - https://research.nhgri.nih.gov/ accessed on 13 May 2016). It is, therefore, critical to investigate any potential MLPA positive result involving a single exon or probe with a second method such as the gene specific confirmation kits or Sanger sequencing to confirm the presence or absence of a true larger deletion or duplication.

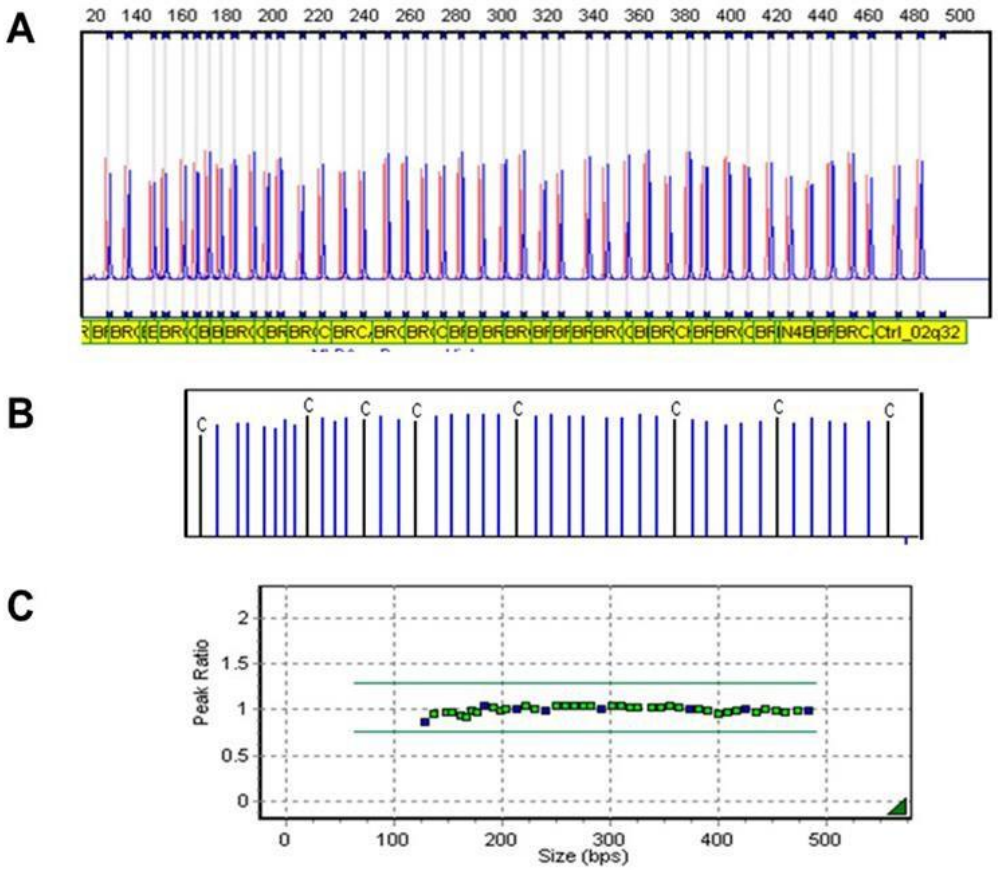

Figure 3.10 Negative MLPA results obtained for *BRCA1* SA BC patient CAM2711. **A.** Raw data for CAM2711 using the SALSA® MLPA® P002-D1 version of the kit. **B.** Dosage histogram representing inter-sample normalization indicating the absence of any duplications or deletions. **C**. Graphical representation of the results obtained for CAM2711, or confirming negative test results for *BRCA1*, as all the probes have a DQ value between 0.7 and 1.3.

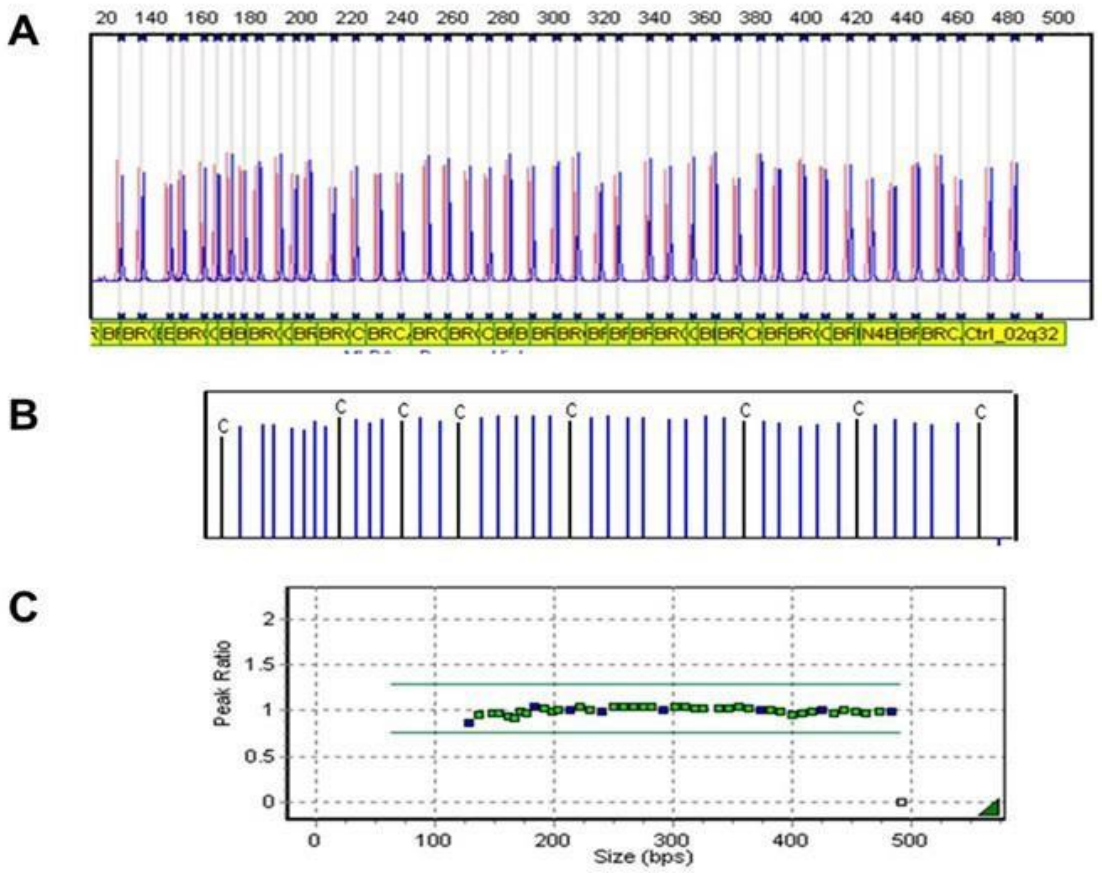

Figure 3.11 Negative MLPA results obtained for *BRCA2* for SA BC patient CAM2711. **A.** Raw data for CAM2711 using the SALSA® MLPA® P045-B3 version of the kit. **B.** Dosage histogram representing inter-sample normalization indicating the absence of any duplications or deletions. **C**. Graphical representation of the results obtained for CAM2711, confirming a negative test result for *BRCA2*, as all the probes have a DQ value between 0.7 and 1.3.

3.3.1 Effect of a five base pair deletion within *BRCA2* exon 9

The *BRCA2* c.771_775delTCAAA (rs80359675) mutation was first described for the Icelandic population by Tulinius *et al.* in 2002. The mutation later proved to be a founder mutation that accounts for 7 - 8% of BC in Iceland (Tulinius *et al.*, 2002; Mikaelsdottir *et al.*, 2004). This mutation was initially referred to as *BRCA2* 999del5, using the older terminology. The deletion results in the creation of a premature truncation of the associated protein, resulting in an ineffective protein with a total mass of 35 kDa (Mikaelsdottir *et al.*, 2004). Since its discovery, this Class V pathogenic mutation has been identified in BC patients throughout the world. It was discovered in Western-, Eastern, and Central Europe, in native Americans, Latin America and now also in an SA patient (BIC - https://research.nhgri.nih.gov/ accessed on 15 May 2016).

BRCA2 c.771_775delTCAAA,p.Asn257_Arg259?fs (rs80359675) has been identified within a single SA Black patient, namely CAM2768. The mutation was initially detected by HRMA and confirmed with DNA Sanger sequencing (Fig. 3.12). The mutation is situated in the middle of the ligation site between the two probes (probe set 01603-L01185, Table 2.7) (Fig 3.12 A). The deleted five bp will affect the hybridisation of the “red” probe the most when it tries to anneal to a partially available target. As the last four bp of the probe cannot bind, the 3’ end of the probe will be unattached. This will affect the ligation of the probes as a minimum of four or potentially five bp (first base of the “blue” probe) are absent. As the MLPA technique is based on immediately adjacent target sequences that need to be ligated into one in order to be amplified, it is proposed that the location of this five bp deletion within exon 9 will cause ligation to fail due to the presence of unattached oligonucleotides. The end result will therefore be a reduced probe signal (false positive), reflecting only the wild type copy of exon 9 (Fig. 3.13 A – C).

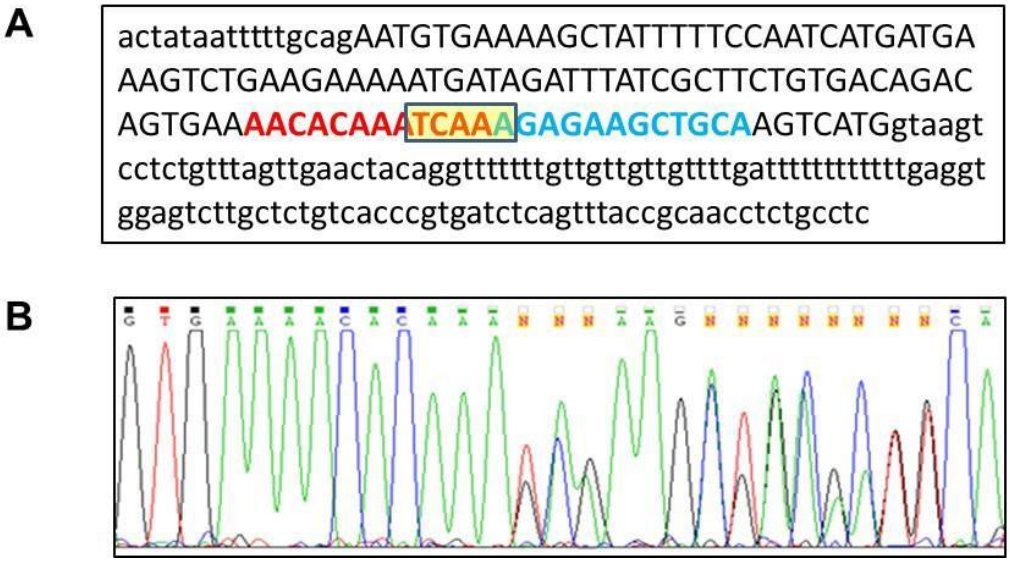

Figure 3.12 Location of the *BRCA2* c.771_775delTCAAA (rs80359675) deletion detected for CAM2768 within exon 9. **A.** Coding sequence for exon 9 (BIC - https://research.nhgri.nih.gov/ accessed on 14 May 2016) with the partial sequence of the two probes utilized within SALSA® MLPA® P045-B3 highlighted in “red” and “blue” respectively. The position of the five bp deletion is indicated by the yellow box. **B.** Electropherogram for *BRCA2* exon 9, indicating the presence of a shift in the reading frame due to the deletion of five bp on the one chromosome.

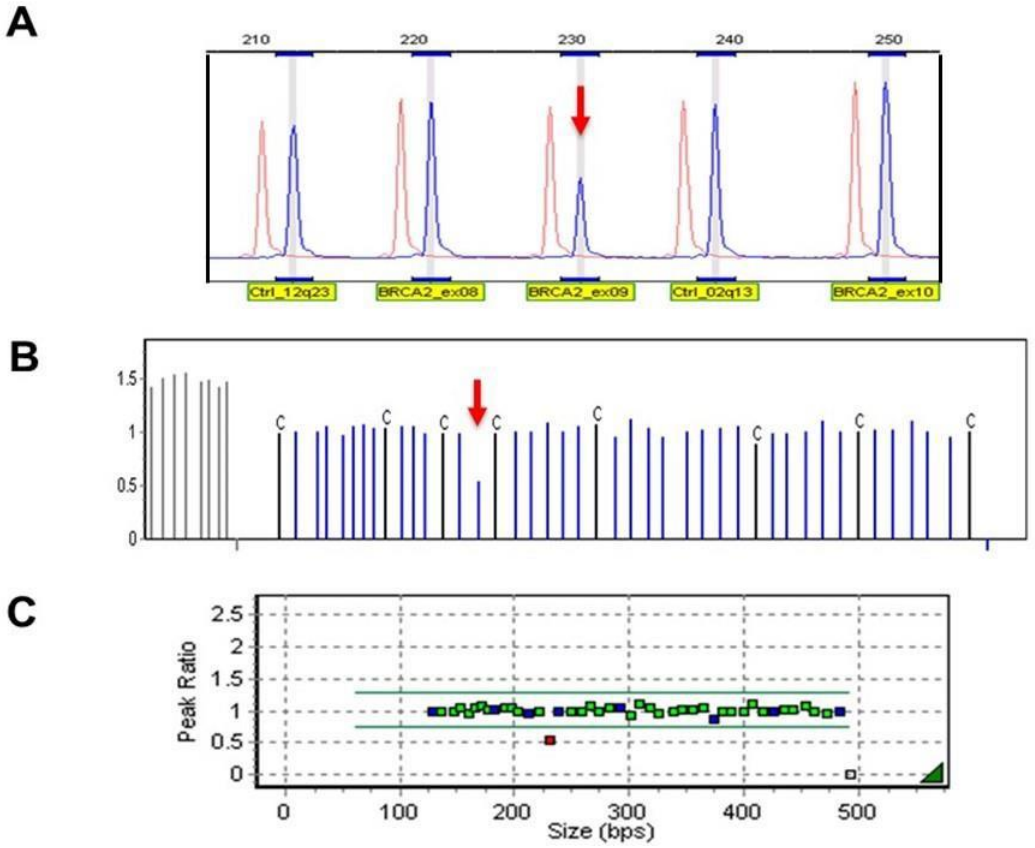

Figure 3.13 Influence of the *BRCA2* c.771_775delTCAAA (rs80359675) deletion on MLPA results, detected for CAM2768. **A.** Raw data for SALSA® MLPA® P045-B3 indicating the deletion of a single probe (indicated by the red arrow) representing exon 20 after intra-sample normalization. **B.** Dosage histogram representing inter-sample normalization indicating the deletion of a single probe (indicated by the red arrow). **C**. Graphical representation of the results obtained for CAM2768, with a single probe having a DQ value below 0.7.

3.3.2 Effect of a two base pair deletion within *BRCA1* exon 20

The *BRCA1* c.5229_5230delAA,p.Gly1743_Arg1744GlyLysfs (rs80357852) mutation was up to now, only reported twice before. The initial discovery was reported by Guy's Hospital in England for an SA family during 1997. The second discovery was submitted to ClinVar (www.ncbi.nlm.nih.gov/clinvar/docs/clinsig/ accessed on 21 May 2016) by Invitae, although the ethnicity of the family or patient is not mentioned. The deletion results in the creation of a premature truncation of the associated protein at amino acid 1828 (BIC). This mutation has been classified as a Class V pathogenic mutation by both the BIC and ClinVar.

BRCA1 c.5229_5230delAA,p.Gly1743_Arg1744GlyLysfs (rs80357852) has been identified within a single SA patient of British/Indian heritage, namely CAM2459. The mutation has been detected by HRMA and confirmed with DNA Sanger sequencing (Fig. 3.14). The location of the mutation is situated in the middle of the ligation site between the two probes (probe set 00783-L00356, Fig 3.14 A). The deleted two bp will affect the hybridisation of the "red" probe the most, as only a partial target area will be available. As the last two bp of the probe cannot bind, the 3' end of the probe (although shorter than for the five bp deletion detected within *BRCA2*) will be unattached. This will affect the ligation of the probes as two bp are absent. The hybridization of the "blue" probe to the target will not be affected. Ligation will again be affected due to the presence of unattached oligonucleotides (loss of two bp on the one chromosome). The end result will, therefore, be a reduced probe signal (false positive), reflecting only the wild type copy of exon 20 (Fig. 3.15 A – C).

3.3.3 Single base pair deletion within *BRCA2* exon 17

The *BRCA2* c.7954_7954delG (rs80359689) mutation was first described for the SA Afrikaner population by Agenbach (2005). The mutation was later proven to be a founder mutation within this specific population group (van der Merwe & van Rensburg, 2009). Although this mutation was predominantly an Afrikaner founder mutation, it was also detected within the Coloured population of the Western Cape

A

tatgacgtgtctgctccacttccattgaaggaagcttctctttctcttatc
ctgatgggttgtgtttggtttctttcagCATGATTTTGAAGTCAGA
GGAGATGTGGTCAATGGAAGAAACCACCAAGGTCCA
AAGCGAGCAAGAGAATCCCAGGACAGAAAGgtaaagct
ccctccctcaagttgacaaaaatctcaccccaccactct**gtattccactc**
ccctttgcagagatggg***ccgcttcattttgtaagact***tattacatacat

B

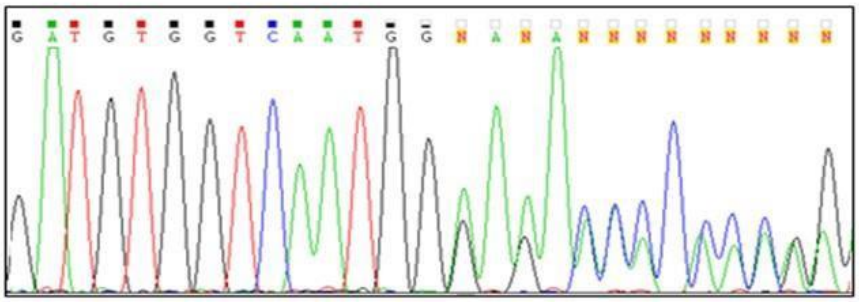

Figure 3.14 Location of the *BRCA1* c.5229_5230delAA (rs80357852) deletion detected for CAM2459 within exon 20. **A.** Coding sequence for exon 20 (BIC - https://research.nhgri.nih.gov/ accessed on 14 May 2016) with the partial sequence of the two probes utilized within SALSA® MLPA® P002-D1 highlighted in red and blue respectively. The position of the two bp deletions are indicated by the yellow box. **B.** Electropherogram for *BRCA1* exon 20, indicating the presence of a shift in the reading frame due to the deletion of two bp on the one chromosome.

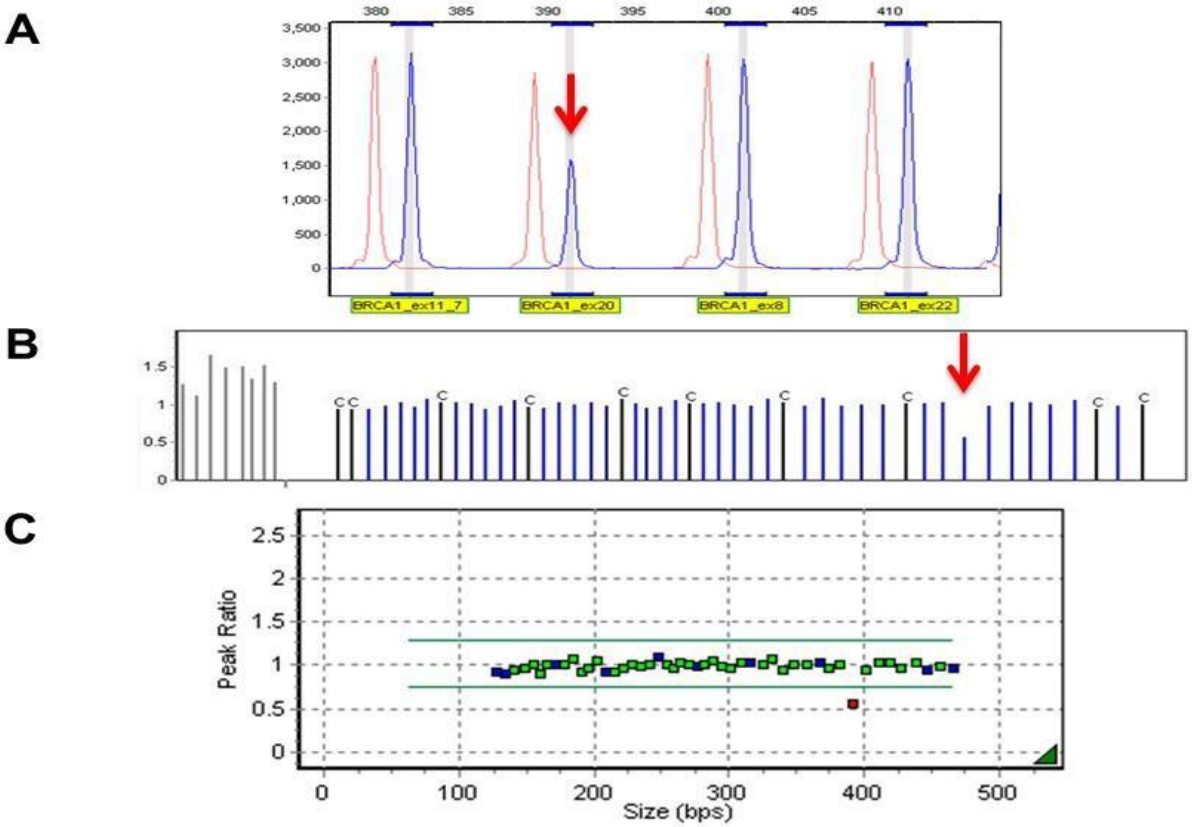

Figure 3.15 Influence of the *BRCA1* c.5229_5230delAA (rs80357852) deletion on MLPA results, detected for CAM2459. **A.** Raw data for SALSA® MLPA® P002-D1 indicating the deletion of a single probe (indicated by the red arrow) representing exon 20 after intra-sample normalization. **B.** Dosage histogram representing inter-sample normalization indicating the deletion of a single probe (indicated by the red arrow). **C**. Graphical representation of the results obtained for CAM2459, with a single probe having a DQ value below 0.7.

(van der Merwe *et al.*, 2012). This mutation is currently the most common founder mutation detected for the Afrikaner and explains the presence of familial BC within the majority of high risk families (van der Merwe & van Rensburg, 2009). This mutation was initially referred to as *BRCA2* 8162delG, using the older terminology. The deletion (*BRCA2* c.7954_7954delG,p.Arg2645Asnfs) results in the creation of premature truncation at amino acid 2647. This Class V pathogenic mutation is restricted to SA population groups and has only been detected within immigrants of SA that are currently living abroad (BIC - https://research.nhgri.nih.gov/ accessed on 20 May 2016).

This SA mutation has been identified in both Caucasian and Coloured patients affected with BC. The mutation was initially detected using single stranded conformational analysis (SSCP) and confirmed with DNA Sanger sequencing (Fig. 3.16). The location of the mutation is not like the other mutations situated in the middle of the ligation site, it is present within the binding site of one of the probes (Fig 3.16 A). The deleted single bp will affect therefore only the hybridisation of the "red" probe as the target area for the "blue" probe is not affected (Fig. 3.16 A). As the deleted base is not at the precise ligation site, ligation will depend on the binding of the three adjacent bp of the "red" probe, in order for ligation to occur (Fig. 3.16 A). Should these bp manage to hybridize, ligation speed will be decreased as the mismatch is closer than five nucleotides of the ligation site (MLPA® DNA Protocol version MDP-005; last revised on 22 Sept 2014). This will result in the destabilizing of the binding of the probe to the target DNA and eventually also result in a decrease in the signal for that probe (Fig. 3.17 A – C).

3.4 MLPA results for SA patients screened

A total of 129 SA BC patients, who previously tested negative for the presence of pathogenic mutations within the familial BC genes, was screened for the presence of larger genomic rearrangements using MLPA (Tables 2.1, 2.2 and 2.3). Of these, 58 were Black, 42 were SA Indian and 29 were Coloured. The age at onset for these BC patients ranged from 21 to 71, with the majority having a positive family history of at least one other BC affected within the immediate family (Tables 2.1, 2.2 and 2.3).

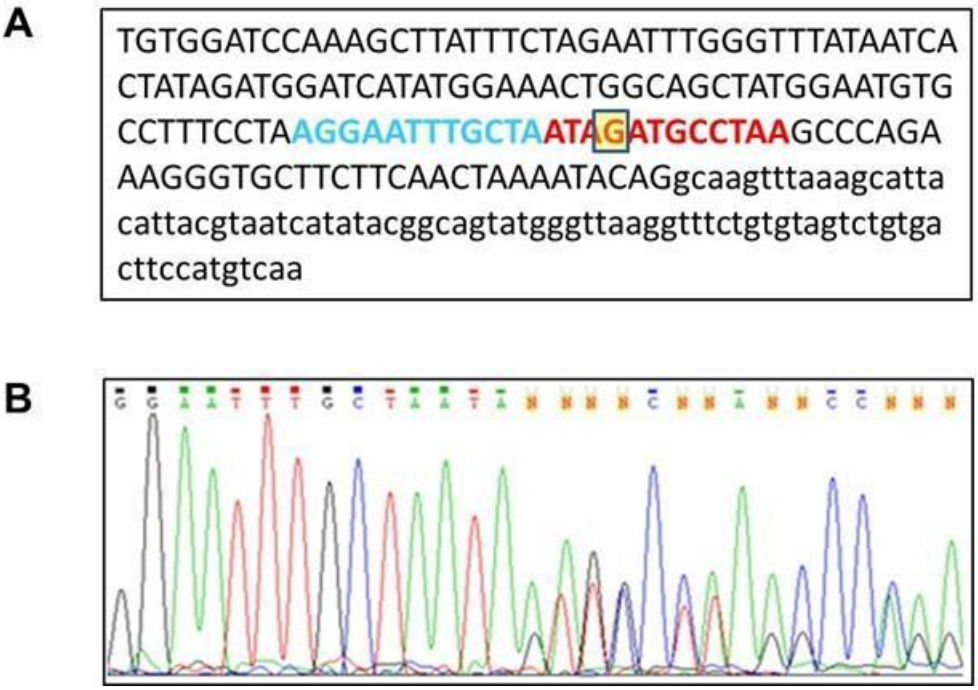

Figure 3.16 Location of the *BRCA2* c.7954_7954delG (rs80359689) deletion detected for CAM2017 within exon 17. **A.** Coding sequence for exon 17 (BIC - https://research.nhgri.nih.gov/ accessed on 20 May 2016) with the partial sequence of the two probes (in reverse) utilized within SALSA® MLPA® P045-B3 highlighted in red and blue respectively. The position of the single bp deletion is indicated by the yellow box. **B.** Electropherogram for *BRCA2* exon 17, indicating the presence of a shift in the reading frame due to the deletion of a single bp on the one chromosome.

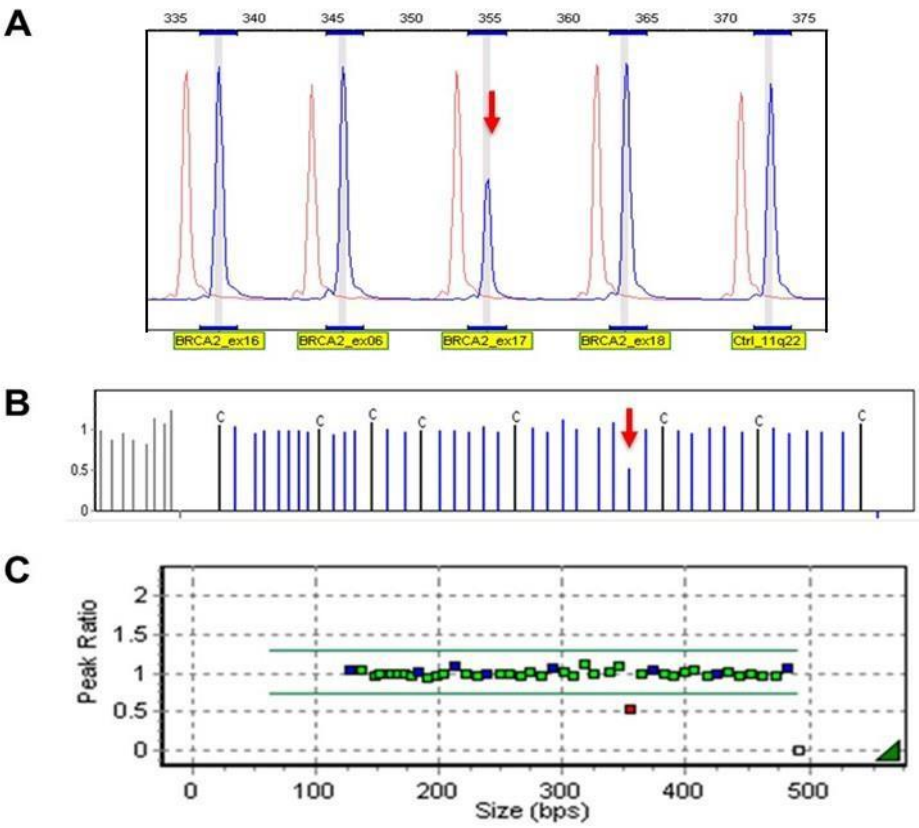

Figure 3.17 Influence of the *BRCA2* c.7954_7954delG (rs80359689) deletion on MLPA results, detected for CAM2017. **A.** Raw data for SALSA® MLPA® P045-B3 indicating the deletion of a single probe (indicated by the red arrow) representing exon 17 after intra-sample normalization. **B.** Dosage histogram representing inter-sample normalization indicating the deletion of a single probe (indicated by the red arrow). **C**. Graphical representation of the results obtained for CAM2017, with a single probe having a DQ value below 0.7.

All these patients were referred from either Tygerberg Hospital or Inkosi Albert Luthuli Hospital in Durban (permission letters – Appendix E & F) and represented patients with either an early age at onset (<45), diagnosed with aggressive disease or who has a positive family history of the disease.

Familial BC is, according to the literature, responsible for 5 – 9% of all BC cases (Ford & Easton, 1995). It was therefore expected that patients representing high risk BC families such as CAM2750, CAM2767, CAM2782, CAM2933 and many others (Table 2.1, 2.2 and 2.3) should carry a pathogenic mutation within either one of the *BRCA* genes. As all these patients tested negative for the presence of smaller deletions/duplications or other deleterious mutations, it was postulated that they might carry larger genomic rearrangements previously not tested for. As the mutation screening approach used was based on the use of qualitative PCR-based techniques that are incapable of detecting this type of mutation, they might have gone unnoticed. This is especially the case for *BRCA1*, which is rich in *Alu* sequences (41.5%) that are known to mediate the occurrence of these type of rearrangements (Smith *et al.*, 1996).

MLPA analysis for 129 SA BC patients ranging from low to high risk revealed the presence of not even a single large rearrangement. All the samples tested negative with the probes for both genes exhibiting DQ values between 0.7 and 1.3 (Figs. 3.10 & 3.11). Patients for which borderline results were observed (Fig. 3.18) were investigated by either repeating the MLPA procedure for that specific gene or by Sanger sequencing. In all of the cases, these probes were either single probes which had a DQ close to 0.7 or 1.3 or were adjacent probes in the probemix that represented non-concurrent exons of the specific gene (Fig. 3.18).

The results obtained for the present study correspond to that of Sluiter and van Rensburg (2010), van Rensburg *et al.* (2014) and Francies *et al.* (2015). Sluiter and van Rensburg were the first to investigate the existence of larger rearrangements within the SA population. They screened 66 SA breast and ovarian cancer patients for the presence of these larger rearrangements within *BRCA1* and *BRCA2*, which included 36 Afrikaner patients. Their study revealed the presence of a single

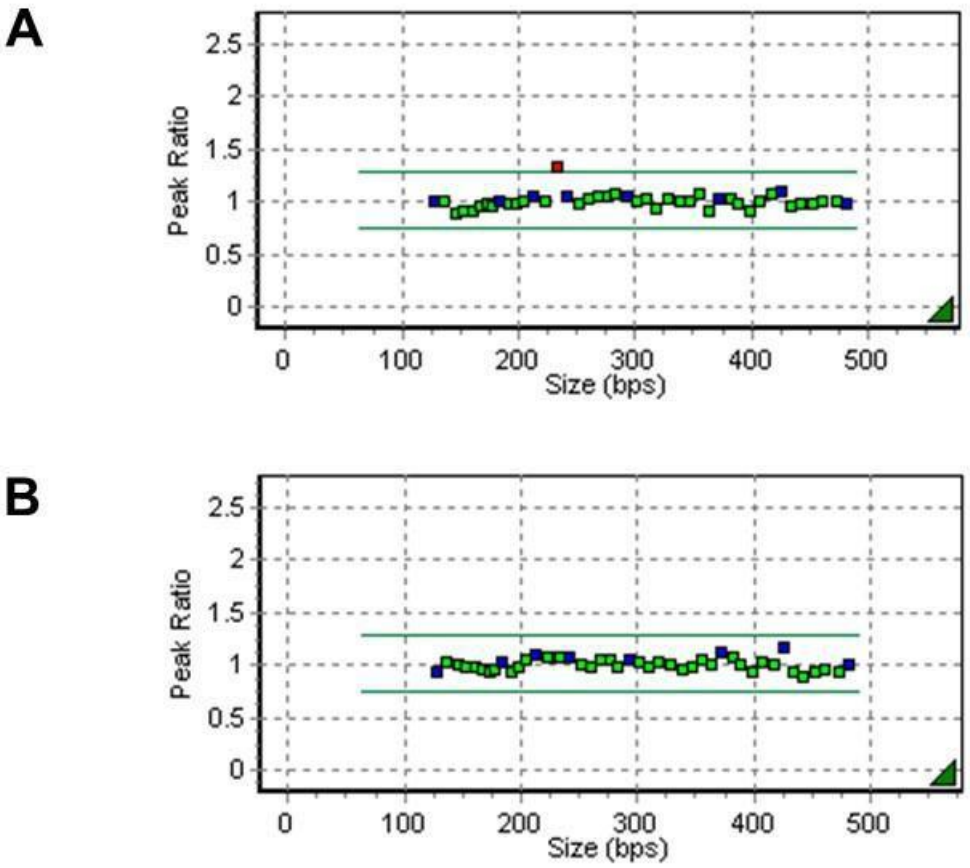

Figure 3.18 Borderline results obtained for CAM2941 using the *BRCA2* SALSA® MLPA® P045-B3 kit. **A.** Graphical representation of the initial MLPA experiment performed for CAM2941 indicating *BRCA2* probe 01603-L01185 (representing exon 9, 229 nt) with a DQ value of ≥1.3 after intra-sample normalization. **B.** Graphical representation of the repeated MLPA experiment performed for CAM2941. *BRCA2* probe 01603-L01185 (representing exon 9, 229 nt) had a DQ value of between 0.7 and 1.3 after intra-sample normalization.

rearrangement in an SA family with a Greek heritage (1.5%). This study did not include patients from the Black, SA Indian or Coloured populations.

The authors continued to search for the presence of larger rearrangements by studying 280 unselected Black BC patients diagnosed at a mean age of 49.51 years (van Rensburg *et al.*, 2014). They detected a novel *BRCA1* deletion which involved the deletion of both exons 21 and 22 (*BRCA1* c.5278-3040_5407+1417del). This mutation was detected within two BC patients, a total of 1% of the samples. The two patients represented different SA tribes, the first was an Nguni-speaking Ndebele woman, and the second was from the Sotho/Tswana-speaking population (van Rensburg *et al.*, 2014). The authors performed haplotype analysis for this specific mutation using five microsatellite markers that flanked/spanned *BRCA1*. Haplotype analysis revealed a common ancestor for this specific mutation.

Based on the results of their study, more Black patients were included in the current study. Although this study did not include any Sotho/Tswana patients, there existed a possibility that the current study might reveal the presence of a BC patient carrying a larger genomic rearrangement, and more specifically the Black African *BRCA1* c.5278-3040_5407+1417del mutation. However, no LGRs were observed from the study. The BC patients of the current study were mainly Zulu and Xhosa residing mostly in Kwazulu-Natal and the Western Cape. As the study participants for the van Rensburg *et al.* (2014) study were collected from the cosmopolitan region of Gauteng, it could be that although the one woman who carried the *BRCA1* c.5278-3040_5407+1417del mutation spoke Sotho/Tswana, she actually represented a different heritage such as the Ndebele. This could explain why the apparent recurrent mutation was not detected in the current study.

The *BRCA1* c.5278-3040_5407+1417del mutation was only the second large rearrangement within *BRCA1* to be reported for African populations. The first was a novel deletion of exon 21 only, that was detected in 0.3% of Nigerian BC cases (Zhang *et al.*, 2010).

The fourth SA study to screen BC patients for the presence of larger genomic rearrangements was performed by Francies *et al.* (2015). The study consisted of a

total sample of 85 Black (78.7%), 16 Caucasian (14.8%), a small number of SA Indian (5 - 4.6%) and Coloured (2 - 1.9%) BC patients, all selected for their triple negative disease. This study also reported no large rearrangements for either of the genes. As patient selection was specifically based on triple negative disease (which is specific for *BRCA1* disease) or an onset of BC before age 40, the involvement of *BRCA1* was expected, but none was found.

Based on the results of the four SA studies, it can be concluded that large rearrangements within the familial BC genes play a negligible role in the disease in SA (Sluiter & van Rensburg, 2010; van Rensburg *et al.* 2014; Francies *et al.*, 2015). Although only two of the studies (Sluiter & van Rensburg, 2010 and the current study) included BC patients based on various selection criteria, a total of 583 SA BC patients have been screened using MLPA. Of this number, only two patients tested positive for the presence of such a mutation (0.34%).

This is in accordance with the trend internationally, as there is a difference in the role that these type of rearrangements play in different ethnic groups and populations. Their contribution ranges from being totally absent (0% in the French Canadian [Moisan *et al.*, 2006] and Ashkenazi Jewish [Stadler *et al.*, 2010] populations), to very low in the Finnish [Pylkas *et al.*, 2008] and SA population) in a specific population to representing founder mutations in others (>50% of all *BRCA1* large genomic rearrangements in Germany [Judkins *et al.*, 2012]; 57.9% in the Portuguese population – [Peixoto *et al.*, 2011]. Founder mutations are therefore responsible for the high rates of large genomic rearrangements in populations such as the Valencian community residing in Eastern Spain (Fachal *et al*., 2014) and the Netherlands population in the Netherlands (Hogervorst *et al.*, 2003).

The majority of these types of founder mutations present in *BRCA1* is due to the *Alu* repetitive DNA elements that contribute to genomic instability due to unequal homologous recombination. This unequal recombination frequently results in deletions and short insertions at the site of the deletion (Hastings *et al.*, 2009). As the majority of large genomic rearrangements are reported for *BRCA1* only, it indicates that *BRCA2* is rarely involved in this type of mutations due to a considerable lower percentage of *Alu* repeats present in the gene. Up to now, only

a single large genomic rearrangement has been identified for *BRCA2*, namely *BRCA2* c.156_157insAluYa5 (NG_012772:g.8686_8687ins AluYa5). This mutation has been detected for the Portuguese population and is the first in this gene to involve an *Alu* element (Machado *et al.*, 2007).

Chapter 4

Conclusion

Diagnostic screening for mutations in *BRCA1* and *BRCA2* is a well-established part of the clinical assessment of familial BC risk. Once a familial pathogenic mutation is detected, pre-symptomatic or susceptibility testing can be offered to unaffected family members to accurately assess their lifetime risks of breast and ovarian cancer. This information then forms the basis for rational decisions regarding uptake of risk management strategies involving regular screening, medical prevention or more radical prophylactic surgery. The effectiveness of this approach is dependent on being able to detect the presence of pathogenic mutations accurately and rapidly.

The frustration and anxiety for the familial BC patients during the time consuming mutation screening process has urged scientists to develop more rapid screening techniques, such as MLPA and next generation sequencing (NGS). The introduction of NGS has revolutionized genomic analysis by dramatically decreasing the cost of sequencing while increasing the throughput. This resulted in a dramatic decrease in the turn-around times for final results.

This study focused on the optimization, validation and the implementation of MLPA as a LGR mutation screening technique. Internationally, the results obtained using this technique have proven its clinical value and place in the mutation screening process of the familial BC genes. The study would have been in vain, if MLPA was to be replaced by NGS. Although NGS is extremely powerful and cost effective, the technique struggles to detect all types of LGRs (discussion at an international symposium on "Optimising the *BRCA* testing pathway for patients with ovarian cancer" held in London, 2015; Walsh *et al.*, 2010; Feliubadalo *et al.*, 2013; Kwong *et al.*, 2015). This is due to the fact that the detection of LGRs by NGS is dependent on the library preparation method and choosing the appropriate bioinformatics

pipeline. In contrast, copy number variation detection by MLPA is based on the hybridization capture method of target enrichment, which may be more challenging for NGS library preparation by amplicon generation (Walsh *et al.*, 2010; Feliubadalo *et al.*, 2013). Based on these principles, MLPA will still play a role in the mutation screening process of BC patients for the familial BC genes, as it cannot as yet be replaced by the NGS method.

This study has resulted in the optimization and validation of the technique in order to be implemented on the diagnostic platform. It is a cost effective and sensitive technique for the detection of LGRs in the familial BC genes. The procedure is uncomplicated and is easy to perform once validated. Although no patients were identified that carried this type of LGRs, valuable MLPA data was generated for the SA population, using five different probe sets (two screening probe sets and one confirmation set for *BRCA1* and two for *BRCA2*, of which one represented the confirmation set). The data corresponded to that of previous studies and supported the tentative hypothesis that LGRs do not seem to play a significant role within the various SA populations. In conclusion, LGRS do not contribute significantly to the familial BC risk within SA.

The data furthermore highlighted the effect that small deletions or duplications might have when situated in the critical ligation site of the probe set. The presence of these smaller mutations could deliver false positive results. The results of this study serve as a warning to pathology laboratories in SA, as the most common Afrikaner founder mutation *BRCA2* c.7954_7954delG (rs80359689) detected in the country, is situated in one of the critical ligation sites of the *BRCA2* SALSA® MLPA® P0045 probemix. The presence of this single base deletion could result in the probes for *BRCA2* exon 17 not ligating, resulting in a decrease in the probe signal and therefore a false positive. This placed emphasis on the confirmation of all potential positive results by using an alternative method or different probe set in order to prevent reporting a false positive result.

In conclusion, carriers of *BRCA* gene mutations are predisposed to a high risk of breast and other cancer types. Full gene sequencing either by conventional

strategies or NGS would give details of the small nucleotide variations, such as point mutations. Techniques such as MLPA would supplement the missing information and provide a more sensitive and accurate detection of LGRs. These complementary approaches are essential for the comprehensive analysis of *BRCA* gene mutations in the future.

Chapter 5

References

Anderson B, Cazap E, Saghir NSEI, Yip C, Khaled H and Otero I, *et al.* (2011). Optimization of breast cancer management in low-resource and middle-resource countries: executive summary if the breast health global initiative consensus, 2010. *Lancet* 12: 387-398.

Armour JA, Barton DE, Cockburn DJ and Taylor GR (2002). The detection of large deletions or duplications in genomic DNA. *Human Mutation* 20: 325-337.

Batzer MA and Deininger PL (2002). *Alu* repeats and human genomic diversity. *Nature Reviews* 3: 370-380.

Belongianni I, Apessos A, Mihalatos M, Razi E, Labropoulos S and Petounis A, *et al.* (2004). Characterization of a novel large deletion and single point mutation in the *BRCA1* gene in a Greek cohort of families with suspected hereditary breast cancer. *Biomedical Central* 24: 61-68.

Brodie SG and Deng C (2001). *BRCA1*-associated tumorigenesis: what have we learned from knockout mice? *Trends in Genetics* 10: s18-s22.

Brown MA, Xu CF, Nicolai H, Griffiths B, Chambers JA and Black D, *et al* (1996). The 5' end of the *BRCA1* gene lies within a duplicated region of human chromosome 17q21. *Oncogene* 12: 2507-2513.

Brzovic PS, Meza JE, King M and Klevit RE (2001). *BRCA1* RING domain cancer-predisposing mutations. *The Journal of Biological Chemistry* 44: 41399-41406.

Bunyan DJ, Eccles DM, Silibourne J, Wilkins E, Thomas SN and Shea-Simonds J, *et al* (2004). Dosage analysis of cancer predisposition genes by multiplex ligation-dependent probe amplification. *British Journal of Cancer* 91: 1155-1159.

Burke W and Austin A (2002). Genetic risk in context calculating the penetrance of *BRCA1* and *BRCA2* mutations. *Journal of the National Cancer Institute* 16: 1185-1187.

Cipollini G, Tommasi S, Paradiso A, Aretini F, Bonatti F and Brunetti I, *et al.* (2004). Genetic alteration in hereditary breast cancer. *Annals of Oncology* 15: i7-i13.

Chen Y, Chen C, Riley J, Alfred D, Chen P and Hoff D, *et al.* (1995). Aberrant subcellular localization of BRCA1 in breast cancer. *Science* 270: 789-789.

Chen C, Li S, Chen Y, Chen P, Sharp D and Lee W (1996). The nuclear localization sequence of the *BRCA1* protein interact with the importation-α subunit of the nuclear transport signal receptor. *The Journal of Biological Chemistry* 61: 32863-32868.

Cherbal F, Bakour R, Adane S, Boualga K, Benais-Pont G and Maillet P (2010). *BRCA1* and *BRCA2* germline mutations screening in Algeria breast/ovarian cancer families. *Disease Marker* 28: 377-384.

Chimusa R, Mbiyavanga M, Mazandu GK, and Mulder NJ (2015). Post Genome-wide Association Study Method for Interaction, Pathway, and Ancestry Analysis in Homogeneous and Admixed Populations. *Bioinformatics,* btv619.

Claus EB, Schildkraut JM, Thompson WD and Risch NJ (1996). The genetic attributable to risk of breast and ovarian cancer. *American Cancer Society* 77: 2318-2324.

Dean G (1963). The prevalence of Porphyria Variegata and disseminated sclerosis in South Africa. *Leech* 33: 114.

Deininger PL and Batzer MA (1999). *Alu* repeats and human disease. *Molecular Genetics and Metabolism* 67: 183-193.

De Jong MM, Nolte IM, te Meerman GJ, van der Graaf WT, Oosterwijk JC and Kleibeuker JH, *et al.* (2002). Genes other than *BRCA1* and *BRCA2* involved in breast cancer susceptibility. *Journal of Medical Genetics* 39: 225-242.

Deng C and Brodie S (2000). Roles of *BRCA1* and its interacting protein. *Biological Essays* 22: 728-737.

Deng C (2002). Tumor formation in *BRCA1* conditional mutant mice. *Environmental and Molecular Mutagenesis* 39: 171-177.

Deng C and Brodie S (2000). Roles of BRCA1 and its interacting proteins. *Biological Essays* 22:728-737.

Del Valle J, Feliubadalo L, Nadal M, Teule A, Miro R, Cuesta R *et al*. (2010). Identification and comprehensive characterization of large genomic rearrangements in the *BRCA1* and *BRCA2* genes. *Breast Cancer Research Treatment* 122: 733–743.

Diamond TM, Sutphen R; Tabano M and Fiorica J (1998). Inherited susceptibility to breast and ovarian cancer. *Current Opinion in Obstetrics & Gynecology* 10: 3-8.

Easton FD, Ford D, Bishop T and Breast Cancer Linkage Consortium (1995). Breast and ovarian cancer incidence in *BRCA1*-mutation carriers. *American Journal of Human Genetics* 56: 265-271.

Ewald IP, Ribeiro PLI, Palmero EI, Cossio SL, Giugliani R and Ashton-Prolla P (2009). Genomic rearrangement in *BRCA1* and *BRCA2*: a literature review. *Genetic and Molecular Biology* 3: 437-446.

Fachal L, Blanco A, Santamarina M, Carracedo A, and Vega A (2014). Large genomic rearrangements of *BRCA1* and *BRCA2* among patients referred for genetic analysis in Galicia (NW Spain): delimitation and mechanism of three novel BRCA1 rearrangements. *PloS one* :3 e93306.

Fackenthal JD and Olopade OI (2007). Breast cancer risk associated with *BRCA1* and *BRCA2* in diverse populations. *Nature Reviews* 7: 937-948.

Feliubadalo L, Lopez-Doriga A, Castellsague E, del Valle J, Menendez M and Tornero E, *et al.* (2013). Next-generation sequencing meets genetic diagnostics: development of a comprehensive workflow for the analysis of *BRCA1* and *BRCA2* genes. *European Journal of Human Genetics* 21:864-870.

Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C and Rebelo M, *et al.* (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns on GLOBOCAN 2012. *International Journal of Cancer* 136: 359-386.

Ford D, Easton DF and Peto J (1995). Estimates of the Gene Frequency of BRCAI and Its Contribution to Breast and Ovarian Cancer Incidence. *American Journal of Human Genetics* 57:1457-1462.

Francies FZ, Wainstein T, De Leeneer K, Cairns A, Murdoch M and Nietz S, *et al.* (2015). *BRCA1*, *BRCA2* and *PALB2* mutations and *CHEK2* c. 1100delC in different South African ethnic groups diagnosed with premenopausal and/or triple negative breast cancer. *Biomed Central* 15: 1.

Frank T.S, Deffenbaugh AM, Reid JE, Hulick M, Ward BE and Lingenfelter B, *et al.* (2002). Clinical characteristics of individuals with germline mutations in *BRCA1* and *BRCA2*: analysis of 10,000 individuals. *Journal of Clinical*

Oncology 6: 1480-1490.

Friedman LS, Szabo CI, Ostermeyer EA, Dowd P, Butler L and Park T, *et al.* (1995). Novel inherited mutations and variable expressivity of *BRCA1* alleles, including the founder mutation 185delAG in Ashkenazi Jewish families. *American Journal of Human Genetics* 57: 1284-1297.

Hall D, Wijsman EM, Roos JL, Gogos JA and Karayiorgou M (2002). Extended intermarker linkage disequilibrium in the Afrikaners. *Genome Research* 12: 956-961.

Hastings PJ, Lupski JR, Rosenberg SM and Ira G (2009). Mechanisms of change in gene copy number. *Nature Reviews* 10: 551-564.

Hogervorst F.B, Nederlof P.M, Gille J.J, McElgunn C.J, Grippeling M, and Pruntel R, *et al.* (2003). Large genomic deletions and duplications in the *BRCA1* gene identified by a novel quantitative method. *Cancer Research* 63: 1449-1453.

Hohenstein P and Fodde R (2003). Of mice and women: genotype-phenotype correlation in *BRCA1*. *Human Molecular Genetics* 12: 271-277.

Homig-Holzel C and Savual S (2012). Multiplex ligation-dependent prove amplification (MLPA) in tumor diagnosis and prognostics. *Diagnosis Molecular Pathology* 4: 189-206.

Judkins T, Rosenthal E, Arnell C, Burbidge LA, Geary W and Barrus T, *et al.* (2012). Clinical significance of large rearrangements in *BRCA1* and *BRCA2*. *Cancer*: 5210-5216.

Knudson AG (1971). Mutation and Cancer: Statistical Study of Retinoblastoma. *Proceedings of National Academy of Science* 68: 820-823.

Korenberg JR and Rykowski MC (1988). Human genome organization: *Alu*, lines, and the molecular structure of metaphase chromosome bands. *Cell*, 53: 391-400.

Kwong A, Chen J, Shin VY, Ho J.C, Law FB and Au C.H, *et al.* (2015). The importance of analysis of long-range rearrangement of *BRCA1* and *BRCA2* in genetic diagnosis of familial breast cancer. *Cancer Genetics* 208: 448-454.

Lahti-Domenici J, Rapakko K, Paakkonen K, Allinen M, Nevanlinna H and Kujala M, *et al.* (2001). Exclusion of large deletions and other rearrangements in *BRCA1* and *BRCA2* in Finnish breast and ovarian cancer families. *Cancer*

Genetics and Cytogenetics 129: 120-123.

Lakhani SR, Gusterson BA, Jacquemier J, Sloane JP, Anderson TJ and van der Vijver MJ, *et al.* (2000). The pathology of familial breast cancer: Histological features of cancer families not attributable to mutations in *BRCA1* and *BRCA2*. *Clinical Cancer Research* 6: 782-789.

Laurila E, Syrjakoski K, Holli K, Kallioniemi A and Karhu R (2005). Search for large genomic alterations of the *BRCA1* gene in Finnish population. *Cancer Genetic and Cytogenetics* 163: 57-61.

Lehur PA, Madarnas P, Devroede G, Perey BJ, Menard DB and Hamade N (1984). Peutz-Jeghers Syndrome: Association of duodenal and bilateral breast cancer in the same patients. *Digestive Disease and Sciences* 29: 178-182.

Lengauer C, Kinzler KW and Volgestein B (1998). Genetic instabilities in human cancers. *Nature* 369: 643-649.

Li FP and Fraumeni JF (1969). Soft-Tissue Sarcomas, Breast Cancer, and Other Neoplasms: A Familial Syndrome? *Annals of Internal Medicine* 71: 747-752.

Liaw D, Marsh DJ, Li J, Dahia PLM, Wang SI and Zheng Z, *et al.* (1997). Germline mutations of the *PTEN* in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nature Genetics* 16: 64-67.

Liu Y and West SC (2002). District functions of *BRCA1* and *BRCA2* in double-strand break repair. *Breast Cancer Research* 4: 9-13.

Lux MP, Fasching PA and Beckman MW (2006). Hereditary breast and ovarian cancer: review and future perspectives. *Journal of Molecular Medicine* 84: 16-28.

Machado PM, Brandão RD, Cavaco BM, Eugénio J, Bento S, Nave M *et al*. (2007). Screening for a BRCA2 rearrangement in high-risk breast/ovarian cancer families: evidence for a founder effect and analysis of the associated phenotypes. *Journal of Clinical Oncology* 25: 2027-2034.

Mayosi B, Flisher A, Lalloo U, Sitas F, Tollman S and Bradshaw D (2009). The burden of non-communicable diseases in South Africa. *Lancet*. 1-14.

Mazoyer S (2005). Genomic rearrangement in the *BRCA1* and *BRCA2* genes. *Human Mutation* 25: 415-422.

Mehheni H, Gehrig C, Nezu J, Oku A, Shimane M and Rossier C, *et al.* (1998). Loss of *LKB1* kinase activity in Peutz-Jeghers syndrome, and evidence for allelic locus heterogeneity. *American Journal of Human Genetics* 63: 1641-1650.

Meijers-Heijboer H, van den Ouweland A, Klijn J, Wasielewski M, de Snoo A and Oldenburg R, *et al.* (2002). Low-penetrance susceptibility to breast cancer due to CHEK2* 1100delC in non-carriers of *BRCA1* or *BRCA2* mutations. *Nature genetics* 31: 55-59.

Mikaelsdottir KE, Valgeirsdottir S, Eyfjord JE and Rafnar T (2004). The Icelandic founder mutation BRCA2 999del5: analysis of expression. *Breast Cancer Research* 4: 284-290.

Miller SA, Dykes DD and Polesky HF (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research* 16: 1215-1215

Miki Y, Swensen J, Shattuck-Eidens D, Futreal A, Harshman K and Tavtigian S, *et al.* (1994). A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* 266: 66-71.

Moisan A, Fortin J, Dumont M, Samson C, Bessette P and Chiquete J, *et al.* (2006). No evidence of *BRCA1/2* genomic rearrangements in high-risk French-Canadian breast/ovarian cancer families. *Genetic Testing* 2: 104-116.

Neuhausen S, Gilweski T, Norton L, Tran T, McGuire P and Swensen J, *et al.* (1996). Recurrent *BRCA2* 6174delT mutations in Ashkenazi Jewish women affected by breast cancer. *Nature Genetics* 13: 126-128.

Neuhausen SL (2000). Founder populations and their uses for breast cancer genetics. *Breast Cancer Research* 2: 77-81.

O'Donovan PJ and Livingston DM (2010). *BRCA1* and *BRCA2*: breast/ovarian cancer susceptibility gene products and participants in DNA double-stranded break repair. *Carcinogenesis* 6: 961-967.

Orbah TI and Olah E (2003). Emerging roles of *BRCA1* alternatively splicing. *Journal of Clinical Pathology* 56: 191-197.

Pamula J, Krzesniak M, Zientek H, Pekala W, Rusin M and Grzywska E (2006). Functional impact of sequences alteration found in *BRCA1* promoter/5'UTR region in breast/ovarian cancer families from upper Silesia Poland. *Hereditary cancer in Clinical Practice* 4(1): 20-24.

Parkin DM, Bray F, Ferlay J and Jemal A (2014). Cancer in Africa 2012. *American Cancer Epidemiology Biomarkers and Prevention* 23: 953-966.

Peioxolo A, Santos C, Pinheiro M, Pinto P, Soares MJ, Rocha P *et al*. (2011). International distribution and age estimation of the Portuguese *BRCA2*

c.156_157insAlu founder mutation. *Breast Cancer Research* 127: 671-679.

Petrij-Bosch A, Peelen T, van Vliet M, van Eijk R, Olmer R and Drusedau M, *et al.* 1997. *BRCA1* genomic deletions are major founder mutations in Dutch breast cancer patients. *Nature Genetics* 17: 341-345.

Petrucelli N, Daly MB, and Feldman GL (2013). *BRCA1* and *BRCA2* hereditary breast and ovarian cancer.

Phelan CM, Kwan E, Jack E, Li S, Morgan C and Aube J, *et al.* (2002). A low frequency of non-founder *BRCA1* mutations in Ashkenazi Jewish breast-ovarian cancer families. *Human Mutation* 20: 352-357.

Pietshmann A, Atri PMM, Hofmann W, Hosseini-Asi SS, Mundlos SS and Peters H (2005). Mutation analysis of *BRCA1* and *BRCA2* genes in Iranian high risk breast cancer families. *Journal of Cancer Research and Clinical Oncology* 131: 552-558.

Puget N, Torchard D, Serova-Sinilnika OM, Lynch HT, Feunteun J and Lenoir GM, *et al.* (1997). A 1-kb *Alu* mediated germ-line deletion removing *BRCA1* exon 17. *American Association of Cancer Research* 57: 828-831.

Pylkäs K, Erkko H, Nikkilä J, Sólyom S and Winqvist R (2008). Analysis of large deletions in BRCA1, BRCA2 and PALB2 genes in Finnish breast and ovarian cancer families. *Bio Medicl Cancer* 8:146-150.

Reeves MD, Yawitch TM, van der Merwe NC, van der Berg HJ, Dreyer G and van Rensburg EJ (2004). *BRCA1* mutations in South African breast/ovarian cancer families: Evidence of a novel founder mutation in Afrikaner families. *International Journal of Cancer* 110: 677-682.

Reeves M (2006). Characterization of *BRCA1* rearrangement in South African breast/ovarian cancer families. (Dissertation).

Roa BB, Boyd AA, Volcik K and Richards CS (1996). Ashkenazi Jewish population frequencies for common mutation in *BRCA1* and *BRCA2*. *Nature Genetics* 14: 185-187.

Roy R, Chun J and Powell S (2012). *BRCA1* and *BRCA2*: different roles in a common pathway of genome protection. *Nature* 12: 68-79.

Schlebusch CM, Dreyer G, Sluiter MD, Yawitch TM, van den Berg HJ and van Rensburg EJ (2010). Cancer prevalence in 129 breast-ovarian families tested for *BRCA1* and *BRCA2* mutations. *South African Journal of Medicine* 100: 113-117.

Schoeman M, Apfflstadt JP, Baatjes K and Urban M (2013). Implementation of a breast cancer genetic service in South Africa – lesson learned. *South African Medical Journal* 8: 529-533.

Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F and Pals G (2002). Relative quantification of 10 nucleic acid sequence by multiplex ligation-dependent probe amplification. *Nucleic Acid Research* 12: e57.

Sigurdson AJ, Hauptmann M, Chatterjee N, Alexander BH, Doody MM and Rutter JL, *et al.* (2004). Kin-Cohort estimates for familial breast cancer risk in relation to variants in DNA base excision repair, *BRCA1* interacting and growth factor genes. *Biomedical Central* 4: 9.

Simon AM, Horwitz AA, Strita LM, Griffin K, Williams RS and Glover JN, *et al.* (2006). *BRCA1* DNA-binding activity is stimulated by *BARD1*. *Cancer Research* 66: 2012-2018.

Sitas F, Madhoo J and Wessie J (1998). Incidence of histologically diagnosed cancer in South Africa, 1993-1995. National Cancer Registry, *South African Institute for Medical Research*, Johannesburg.

Sluiter MD and van Rensburg EJ (2011). Large genomic rearrangements of *BRCA1* and *BRCA2* genes: review of the literature and report of a novel *BRCA1* mutation. *Breast Cancer Research Treatment* 125: 325-349.

Smith TM, Lee MK, Szabo CI, Jerome N, McEuen M and Taylor M, *et al.* (1996). Complete genomic sequence and analysis of 117 kb of human DNA containing the gene *BRCA1*. *Genome Research* 6: 1029-1049.

Sokolenko AP, Mitiushkina NV, Buslov KG, Bit-Sava EM, Iyevleva AG, Chekmariovaa EV *et al*. (2006). High frequency of BRCA1 5382insC mutation in Russian breast cancer patients. *European Journal of Cancer* 42: 1380-1384.

Somdyala N, Bradshaw D, Gelderblom W and Parkin D (2010). Cancer incidence in a rural population of South Africa, 1998–2002. *International Journal of Cancer* 127: 2420-2429.

Spain BH, Larson CJ, Shihabuddin LS, Gege FH and Verma I (1999). Truncated *BRCA2* is cytoplasmic: implications for cancer-linked mutations. *Proceedings of the National Academy of Sciences* 24: 13920-13926.

Stadler ZK, Saloustros E, Hansen NA, Schluger AE, Kauff ND and Offit K, *et al.* (2010). Absence of genomic *BRCA1* and *BRCA2* rearrangements in

Ashkenazi Jewish breast and ovarian cancer families. *Breast Cancer Research Treatment* 123: 581-585.

Stratton MR and Rahman N (2008). The emerging landscape of breast cancer susceptibility. *Nature Genetics* 17: 17-22.

Tavtigian SV, Simard J, Rommens J, Couch F, Shattuck-Eidens D, and Neuhausen *et al.* (1996). The complete *BRCA2* gene and mutations in chromosome 13q-linked kindreds. *Nature Genetics* 12: 333-337.

Thompson D and Easton D (2002). Cancer Incidence in BRCA1 Mutation Carriers. *Journal of the National Cancer Institute* 94: 1358-1365.

Thompson D and Easton D (2004). The Genetic Epidemiology of Breast Cancer Genes. *Journal of Mammary Gland Biology and Neoplasia* 9: 221-236.

Tonin P, Weber B, Offit K, Couch F, Rebbeck TR and Neuhausen S, *et al.* (1996). Frequency of recurrent *BRCA1* and *BRCA2* mutations in Ashkenazi Jewish breast cancer families. *Nature Medicine* 2: 1179-1183.

Torre LA, Bray F, Siegel R, Ferlay J, Lortet-tielent J and Jemal A (2015). Global cancer statics, 2012. *Cancer Journal of Clinicians* 65: 87-108.

Tulinius H, Olafsdottir HG, Sigvaldason H, Arason A, Barkardottir RB, Egilsson V *et al*. (2002). The effect of a single BRCA2 mutation on cancer in Iceland. *Journal of Medical Genetics* 39:457–462.

Tutt A and Ashworth A (2002). The relationship between roles of *BRCA* genes in DNA repair predisposition. *Trends in Molecular Medicine* 12: 571-576.

Unger MA, Nathanson KL, Calzone K, Antin-Ozerkis D, Shih HA *et al*. (2000). Screening for Genomic Rearrangements in Families with Breast and Ovarian Cancer Identifies BRCA1 Mutations Previously Missed by Conformation-Sensitive Gel Electrophoresis or Sequencing. *American Journal of Human Genetics* 67:841–850.

Van der Merwe NC and Van Rensburg EJ (2009). Hereditary breast/ovarian cancer and *BRCA* mutations: a South African perspective. *Current Oncology* 16: 347.

Van der Merwe NC, Hamel N, Schneider S-R, Apffelstaedt JP, Wijnen JT and Foulkes WD (2012). A founder *BRCA2* mutation in non-Afrikaner breast cancer patients of the Western Cape. *Clinical Genetics* 81: 179-184.

Venkitaraman AR (2002). Cancer susceptibility and the functions of *BRCA1* and *BRCA2*. *Cell* 106: 171-182.

Vorobiof DA, Sitas F and Vorobiof G (2001). Breast cancer incidence in South Africa. *Journal of Clinical Oncology* 19: 125-127.

Wagner JE, Tolar J, Levran O, Scholl T, Deffenbaugh A and Satagopan, *et al.* (2004). Germline mutations in *BRCA2*: shared genetic susceptibility to breast cancer, early onset leukemia, and Fanconi anemia. *Blood* 10: 3226-3229.

Walsh T, Lee M.K, Casadei S, Thornton A.M, Stray S.M and Pennil C, *et al.* (2010). Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing. *Proceedings of the National Academy of Sciences* 107: 12629-12633.

Warren M, Smith A, Partridge N, Masabanda J, Griffin D and Ashworth A (2002). Structural analysis of the chicken *BRCA2* gene facilitates identification of functional domains and disease causing mutations. *Human Molecular Genetics* 7: 841-851.

Welcsh PL, Owens KN and King M (2000). Insight into the function of *BRCA1* and *BRCA2*. *Trends in Genetics* 2: 69-74.

Welcsh PL and King M (2001). *BRCA1* and *BRCA2* and the genetics of breast and ovarian cancer. *Human Molecular Genetics* 7: 705-713.

Williams RS, Chasman DI, Hau DD, Hui B, Lau AY and Glover JN (2003). Detection of protein folding defects caused by *BRCA1*-BRCT truncation and missense mutations. *Journal of Biological Chemistry* 52: 53007-53016.

Williams LH, Choong D, Johnson SA and Campbell IG (2006). Genetic and epigenetic analysis of *CHEK2* in sporadic breast, colon, and ovarian cancers. *Clinical Cancer Research* 12: 6967-6972.

Wong AK, Pero R, Ormonde PA, Tavtigian SV and Bartel P (1997). RAD51 interact with the evolutionary conserved BRC motifs in the breast cancer susceptibility gene brca2. *The Journal of Biological Chemistry* 51: 310941-31944.

Wooster R, Neuhausen SL, Mangion J, Quirk Y, Ford D and Collins N, *et al.* (1994). Localization of a breast cancer susceptibility gene, *BRCA2*, to chromosome 13q12-13. *Science* 265: 2088-2090.

Wooster R, Bignll G, Lancaster J, Swift S, Seal S and Mangion J, *et al.* (1995). Identification of the breast cancer susceptibility gene BRCA2. *Nature* 378, 789-792.

Wooster R and Weber BL (2003). Breast and ovarian cancer. *The New England Journal of Medicine* 348: 2339-2347.

Xu C, Brown MA, Chamber JA, Griffiths B, Nicola H and Solomon E (1995). Distinct transcription start sites generate two forms of BRCA1 mRNA. *Human Molecular Genetics* 12: 2259-2264.

Xu C, Chamber JA and Solomon E (1997). Complex regulation of the *BRCA1* gene. *The Journal of Biological Chemistry* 272: 20994-20997.

Yang H, Jeffrey PD, Miller J, Kinnucan E, Sun Y and Thoma NH, *et al.* (2002). *BRCA2* function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. *Science* 297: 1837-1848.

Zhang W, Edwards A, Fan W, Deininger P and Zhang K (2011). *Alu* distribution and mutation types of cancer genes. *Biomedical Genomics* 12: 157.

Zhang B, Beeghly-Fadiel A, Long J and Zheng W (2011b). Genetic variants associated with breast-cancer risk: comprehensive research synopsis, meta-analysis, and epidemiological evidence. *Lancet Oncology* 12:477-488.

Zheng J, Fackenthal J, Huo D, Zheng Y and Olopade O (2010). Searching for large genomic rearrangements of the BRCA1 gene in a Nigerian population. *Breast Cancer Research Treatment* 124:573-577.

List of electronic references

Breast Cancer Information Core (BIC). [Online] Available http://www.ncbi.nlm.nih.gov/gene/672. 2015.

Evol3000: Bottlenecks and founder effect. [Online] Available at http://wallace.genetics.uga.edu/groups/evol3000/wiki/fb221/Bottlenecks_and_Founder_Effects.html. 2015.

Genetic Counselling. [Online] Available at http://www.geneticcounselling.co.za. 2015.

NanoDrop™. [Online] Available at www.nanodrop.com/.../T042-NanoDrop-Spectrophotometers-Nucleic-Acid. 2015.

National Centre for Biotechnology Information (NCBI). [Online] Available at http://www.ncbi.nlm.nih.gov/books/NBK1247/. 2015.

National Human Genome Research Institute (NHGRI). [Online] Available at http://research.nhgri.nih.gov/bic/. 2015.

National Institute for Occupational Health, Division of the National Health Laboratory Services [Online] Available at www.nioh.ac.za/assets/files/NCR_Final_2010_tables(1).pdf

MRC Holland. [Online] Available at http://www.mrc-holland.com/. 2015.

Multiplex Ligation-dependent Probe Amplification (MLPA). [Online] Available at http://www.mlpa.com/WebForms/WebFormMain.aspx?Tag. 2015.

Appendices

Appendix A

CONSENT FOR ADDITIONAL DNA STUDIES

1. I, _____ , request that an attempt be ma using genetic material to assess the probability that: I (name:_____) might have inheritec disease-causing mutation in the gene for: **FAMILIAL BREAST CANCER.**

2. I understand that the genetic material for analysis is to be obtained from: blood cells.

3. I acknowledge that a portion of the sample will be stored indefinitely for later use for:

   (a) possible re-analysis for the same disorder

   (b) research purposes (for any disorder), subject to the approval of the Institution's Ethics Committee, provided that any information from such research will remain confidential and anonymous.

4. The results of the analysis carried out on this sample of stored biological material will be made known to me, via my doct

5. I authorise my doctor(s) to provide relevant clinical details in order for the research to be published.

6. I have been informed that:

   (a) **I will not receive any rumeneration for my involvement in this study.**

   (b) **I will not endure additional financial costs due to the tests being performed.**

   (c) The risks and benefits associated with genetic analysis have been explained to me.

   (b) The analysis procedure is specific to the genetic condition mentioned above and cannot determine the complete genetic makeup of an individual.

   (c) The genetics laboratory is under an obligation to respect and maintain medical confidentiality.

   (d) Genetic analysis may not be informative for some families or family members.

   (e) **where biological material is used for research purposes, there may be no direct benefit to me.**

7. I understand that I may withdraw or modify my consent for any aspect of the above at any time without this affecting my future medical care.

8. **ALL OF THE ABOVE HAS BEEN EXPLAINED TO ME IN A LANGUAGE THAT I UNDERSTAND AND MY QUESTIONS ANSWERE BY:**

9. **Address:**

Tel. no.:

Signature _____ Date: _____
of person giving consent Capacity _____

Patient signature: _____ **Witnessed consent:**_____

Appendix B

UNIVERSITY OF THE FREE STATE
UNIVERSITEIT VAN DIE VRYSTAAT
YUNIVESITHI YA FREISTATA
UFS UV

Research Division
Internal Post Box G40
☎(051) 4052812
Fax (051) 4444359

E-mail address: StraussHS@ufs.ac.za

Ms H Strauss

2012-04-17

DR NC VAN DER MERWE
DIVISION OF HUMAN GENETICS
FACULTY OF HEALTH SCIENCES
UFS

REC Reference nr 230408-011
IRB nr 00006240

Dear Dr van der Merwe

ETOVS NR 65/08
PROJECT TITLE: SCREENING OF YOUNG AFRICAN BREAST CANCER PATIENTS FOR THE PRESENCE OF DELETERIOUS BRCA1 AND BRCA2 MUTATIONS.

- You are hereby kindly informed that the Ethics Committee approved the following at the meeting on 10 April 2012:
  - *Request to extend the study period for three years.*
- Committee guidance documents: Declaration of Helsinki, ICH, GCP and MRC Guidelines on Bio Medical Research. Clinical Trial Guidelines 2000 Department of Health RSA; Ethics in Health Research: Principles Structure and Processes Department of Health RSA 2004; Guidelines for Good Practice in the Conduct of Clinical Trials with Human Participants in South Africa, Second Edition (2006); the Constitution of the Ethics Committee of the Faculty of Health Sciences and the Guidelines of the SA Medicines Control Council as well as Laws and Regulations with regard to the Control of Medicines.
- A progress report should be submitted within one year of approval of long term studies and a final report at completion of both short term and long term studies.
- Kindly refer to the ETOVS/ECUFS reference number in correspondence to the Ethics Committee secretariat.

Yours faithfully

.....
CHAIR: ETHICS COMMITTEE

205 Nelson Mandela Drive/Rylaan, Park West/Parkwes, Bloemfontein 9301, South Africa/Suid-Afrika
P.O. Box/Posbus 339, Bloemfontein 9300, South Africa/Suid-Afrika, T: +27(0)51 401 9111, www.ufs.ac.za

Appendix C

UNIVERSITY OF THE FREE STATE
UNIVERSITEIT VAN DIE VRYSTAAT
YUNIVESITHI YA FREISTATA
UFS UV

Research Division
Internal Post Box G40
☎(051) 4017795
Fax (051) 4444359

E-mail address: EthicsFHS@ufs.ac.za

Ms J du Plessis/gn

2014-11-12

REC Reference nr 230408-011
IRB nr 00006240

MR HMVE COMBRINK
DIVISION OF HUMAN GENETICS
FACULTY OF HEALTH SCIENCES
UFS

Dear Mr Combrink

ECUFS NR 107/2014
PROJECT TITLE: MOLECULAR SCREENING OF THE SOUTH AFRICAN INDIAN POPULATION FOR BRCA1 AND BRCA2 USING HIGH RESOLUTION MELTING ANALYSIS.

1. You are hereby kindly informed that at the meeting on 04 November 2014 the Ethics Committee approved the above project after all conditions were met when the following was submitted:

   - ***Revised Information Leaflet and satisfying answers to the reviewer's questions***

2. Committee guidance documents: Declaration of Helsinki, ICH, GCP and MRC Guidelines on Bio Medical Research. Clinical Trial Guidelines 2000 Department of Health RSA; Ethics in Health Research: Principles Structure and Processes Department of Health RSA 2004; Guidelines for Good Practice in the Conduct of Clinical Trials with Human Participants in South Africa, Second Edition (2006); the Constitution of the Ethics Committee of the Faculty of Health Sciences and the Guidelines of the SA Medicines Control Council as well as Laws and Regulations with regard to the Control of Medicines.

3. Any amendment, extension or other modifications to the protocol must be submitted to the Ethics Committee for approval.

4. The Committee must be informed of any serious adverse event and/or termination of the study.

5. All relevant documents e.g. signed permission letters from the authorities, institutions; changes to the protocol, questionnaires etc. have to be submitted to the Ethics Committee before the study may be conducted (if applicable).

6. A progress report should be submitted within one year of approval of long term studies and a final report at completion of both short term and long term studies.

Appendix D

UNIVERSITY OF THE FREE STATE
UNIVERSITEIT VAN DIE VRYSTAAT
YUNIVESITHI YA FREISTATA
UFS UV

[redacted]

Research Division
Internal Post Box G40
☎(051) 4052812
Fax (051) 4444359

E-mail address: EthicsFHS@ufs.ac.za

Ms J Du Plessis

2014-09-22

REC Reference nr 230408-011
IRB nr 00006240

DR N VAN DER MERWE
DIVISION OF HUMAN GENETICS
FACULTY OF HEALTH SCIENCES
UFS

Dear Dr van der Merwe

ECUFS NR 108/2014
MR J OOSTHUIZEN **DIVISION OF HUMAN GENETICS**
PROJECT TITLE: MOLECULAR SCREENING OF SOUTH AFRICAN COLOURED BREAST CANCER PATIENTS FOR THE PRESENCE OF BRCA MUTATIONS USING HIGH RESOLUTION MELTING ANALYSIS

1. You are hereby kindly informed that at the meeting on 16 September 2014 the Ethics Committee approved the above project after all conditions have been met when the following was submitted:

   - ***Signed permission letter from the Tygerberg Hospital***

2. You are hereby kindly informed that the Ethics Committee is concerned about this study being the same as the study with the project title: *"Molecular screening of the South African Indian population for BRCA1 and BRCA2 using High resolution melting analysis". Kindly note that the* studies may be the same but the write up for the M.Med studies should not be the same word for word.

3. Committee guidance documents: Declaration of Helsinki, ICH, GCP and MRC Guidelines on Bio Medical Research. Clinical Trial Guidelines 2000 Department of Health RSA; Ethics in Health Research: Principles Structure and Processes Department of Health RSA 2004; Guidelines for Good Practice in the Conduct of Clinical Trials with Human Participants in South Africa, Second Edition (2006); the Constitution of the Ethics Committee of the Faculty of Health Sciences and the Guidelines of the SA Medicines Control Council as well as Laws and Regulations with regard to the Control of Medicines.

4. Any amendment, extension or other modifications to the protocol must be submitted to the Ethics Committee for approval.

5. The Committee must be informed of any serious adverse event and/or termination of the study.

6. All relevant documents e.g. signed permission letters from the authorities, institutions, changes to the protocol, questionnaires etc. have to be submitted to the Ethics Committee before the study may be conducted (if applicable).

Appendix E

health

Department:
Health
PROVINCE OF KWAZULU-NATAL

INKOSI ALBERT LUTHULI CENTRAL HOSPITAL
Ethekwini Health District
Office of the Chief Executive Officer
800 Vusi Mzimela (Bellair) Road, Mayville, 4091
Private Bag X03, Mayville, 4058
Tel: 031 - 240 1034
Fax: 031 - 240 1005
Email: Gugu.Duma@ialch.co.za
www.kznhealth.gov.za

Enquiries: Dr S.T. Mtshali
Date: 15 August 2014

Ms Gugu Khumalo
c/o Department of Health KwaZulu-Natal

Email: gugu.khumalo@kznhealth.gov.za

RESEARCH PROPOSAL

A research proposal has been submitted to me by another centre requesting that data of IALCH patients be used for the study "MOLECULAR SCREENING OF THE SOUTH AFRICAN INDIAN POPULATION FOR BRCA1 AND BRCA2 .

The protocol has been forwarded to the Research Committee in Pietermaritzburg for clearance and the research has been approved.

I wish to inform you that I consent to the request for samples of IALCH patients to be used for the research in question.

Yours sincerely,

Dr S.T. Mtshali
Chief Executive Officer

Appendix F

Western Cape Government
Health

Tygerberg Hospital

REFERENCE: Research Projects
ENQUIRIES: Dr G G Marinus
TELEFONE: 021 938-6267

ETHICS NO: ECUF 108/2014

Molecular screening of South African coloured breast cancer patients for the presence of BRCA mutations using high resolution melting analysis

Dear Van der Merwe

PERMISSION TO CONDUCT YOUR RESEARCH AT TYGERBERG HOSPITAL

In accordance with the Provincial Research Policy and Tygerberg Hospital Notice No 40/2009, permission is hereby granted for you to conduct the above-mentioned research here at Tygerberg Hospital.

DR D ERASMUS
CHIEF EXECUTIVE OFFICER
Date: 22 August 2014

Administration Building, Francie van Zijl Avenue, Parow, 7500
tel: +27 21 938-6267 fax: +27 21 938-4890
Private Bag X3, Tygerberg, 7505
www.capegateway.go.v.za

Appendix G

health
Department:
Health
PROVINCE OF KWAZULU-NATAL

Health Research & Knowledge Management sub-component
10 – 103 Natalia Building, 330 Langalibalele Street
Private Bag x9051
Pietermaritzburg
3200
Tel.: 033 – 3953189
Fax.: 033 – 394 3782
Email.: hrkm@kznhealth.gov.za
www.kznhealth.gov.za

Reference : HRKM194/14
Enquiries: Mrs G Khumalo
Telephone : 033 – 395 3189

Dear Dr NC van der Merwe

Subject: Approval of a Research Proposal

1. The research proposal titled **'Molecular screening of the South African Indian population for BRCA1 and BRCA2'** was reviewed by the KwaZulu-Natal Department of Health (KZN-DoH).

The proposal is hereby **approved** for research to be undertaken at Inkosi Albert Luthuli Central Hospital.

2. You are requested to take note of the following:
   a. Make the necessary arrangement with the identified facility before commencing with your research project.
   b. Provide an interim progress report and final report (electronic and hard copies) when your research is complete.

3. Your final report must be posted to **HEALTH RESEARCH AND KNOWLEDGE MANAGEMENT, 10-102, PRIVATE BAG X9051, PIETERMARITZBURG, 3200** and e-mail an electronic copy to hrkm@kznhealth.gov.za

For any additional information please contact Mrs G Khumalo on 033-395 3189.

Yours Sincerely

Dr. E Lutge

Chairperson, KwaZulu-Natal Health Research Committee

Date: 21/09/14

Appendix H

National Health Laboratory Service
Human Genetics \ Mensgenetika

Tel: +27 51 4053047
Fax: +27 51 4441161
Practice/Praktyk no/5200296

Block C
Faculty of Health Sciences
University of the Free State
Bloemfontein

P.O. Box 339 (G11)
UFS
Bloemfontein 9300

Prof M Theron
Head of Department of Human Genetics
National Health Laboratory Services
Faculty of Health Sciences, UFS
BLOEMFONTEIN

13 June 2014

Dear Prof M Theron

Re: Permission for use of laboratory space and blood samples from clinic patients

I am in the process of expanding our current research project on familial breast cancer involving the genes *BRCA1* and *BRCA2*. We aim to screen young breast cancer patients (especially of Indian, Mixed Ancestry or African decent) or patients with a positive family history (any ethnic group). Blood samples from patients will be collected for DNA extraction. The project involves screening these patients for disease-causing mutations present in these genes, with the hope to establish a more informative diagnostic mutation screening protocol for each of the population groups within South Africa. It is currently limited to the Afrikaner population.

I would therefore like to ask your permission to use the space and equipment of the Division of Human Genetics (Molecular Laboratory) for this project. We hope to present the data at the local Faculty Forum as well as on other national congresses, depending on patient numbers and the results obtained. This study will be ongoing and will commence in April/May of 2014, after ethic approval has been obtained.

Your prompt reply will be appreciated.

Yours sincerely

Dr NC vd Merwe PhD
Principal investigator
E-mail: gnmgncy.MD@mail.uovs.ac.za
On behalf of: Pakiso Moeti
HMVE Combrink
J Oosthuizen

Approved :

Division of Human Genetics
University of the Free State &
2014 -06- 13
NHLS Universitas Tertiary Academic
Laboratories
Bloemfontein