Development and application of a pathology supported pharmacogenetic test for improved clinical management of South African patients with breast cancer and associated co-morbidities

BY

NICOLE VAN DER MERWE

Thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Anatomical Pathology in the faculty of Medicine and Health Sciences at Stellenbosch University

Supervisor:
Professor Maritha J Kotze

Co- Supervisors:
Professor Susan Janse van Rensburg,
Professor Juanita Bezuidenhout,
Doctor Fredrieka Pienaar

Division of Anatomical Pathology
Department of Pathology, Faculty of Medicine and Health Sciences
Stellenbosch University

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date.....**March 2016**.....

Summary

Three major challenges in the field of breast cancer have been identified as research priorities for this study. The first is the need to combine genetic testing of high-risk patients with familial breast cancer with pharmacogenetics to reduce recurrence risk in cancer survivors due to drug failure as a consequence of anti-cancer treatment that does not match the patient's genotype. The second is the delineation of key pathways through which genes implicated in breast cancer and associated co-morbidities can serve as nutritional and drug targets across diagnostic boundaries. The third is the discovery of genetic alterations underlying familial breast cancer not attributed to mutations in the two major tumour suppressor genes, BRCA1 and BRCA2.

The study population consisted of 164 breast cancer patients (60 Coloured/Mixed Ancestry and 104 Caucasian), of whom 88 patients were selected from a total of 813 individuals who provided informed consent for inclusion of their data in a genomics database resource generated at the interface between the laboratory and routine clinical practice. In addition, DNA samples of 101 cancer-free individuals above the age of 65 years were available for clinical validation of potentially causative variants in an extended female control group.

In the first phase of this study, real-time polymerase chain reaction (PCR) TaqMan© technology was used to confirm the potential value of adding pharmacogenetic testing (CYP2D6 allele 4) to standard immunohistochemistry (IHC)-based breast tumour subtyping complemented by BRCA mutation screening and/or microarray gene profiling in eligible patients. In phase two of the study, common genetic risk factors for cardiovascular disease (CVD) were shown to be significantly associated with earlier age (10 years on average) of breast cancer onset/diagnosis (APOE E4 allele p=0.003; 95% CI: 4-15) and body mass index (BMI) (MTHFR 1298 A>C; p=0.01; 95% CI: 3-14) in patients stratified according to estrogen receptor (ER) status, after adjustment for potential confounders. Age at diagnosis/onset of breast cancer was significantly lower in patients with ER-negative versus ER-positive tumours, after adjustment for ethnicity (p=0.022), while BMI was significantly higher in patiens with ER-positive compared to ER-negative tumours after adjustment for age, ethnicity, and family history of cancer (p=0.035). These findings contributed to the development of an exome pre-screening algorithm (EPA) used in part 3 of this study to select three genetically uncharacterized breast cancer patients for whole exome sequencing (WES) performed in comparison with three ethnically concordant cancer-free controls. WES followed by variant calling using both the standard human genome reference sequence (hg19) and an ethnically concordant major allele reference genome (MARS) revealed a more than 20% discrepancy in the number of gene variants identified in the same samples. After exclusion of a large number of false-positives caused by minor alleles in

hg19, two rare missense mutations (<1%) were identified in a family with ER-positive breast cancer: RAD50 R385C and MUC1 Q67E. Three different bioinformatics tools were used to predict functionality and both mutations were confirmed by Sanger sequencing and/or real-time PCR in the Pathology Research Facility (PRF) laboratory. Neither the RAD50 nor the MUC1 missense mutation were identified in the exomes of an unrelated breast cancer patient with triple-negative breast cancer or three population-matched control individuals.

This study led to the development of a pathology-supported genetic testing framework for WES beyond the limitations of single-gene BRCA mutation screening in South African breast cancer patients. Our findings support previous WES results indicating that the majority of genetically uncharacterised familial breast cancer may be caused by a combination of low-moderate penetrance mutations exerting their effect in a high-risk environment reflected by high BMI. WES enables identification of genetic risk factors of relevance to both cancer development and tailored therapeutic intervention in a single genetic test.

Opsomming

Drie belangrike uitdagings in die veld van borskanker-navorsing is geïdentifiseer as die fokus vir navorsing in hierdie studie. Die eerste is die nodigheid om genetiese toetsing van hoë-risiko pasiënte met familiële borskanker te kombineer met farmakogenetika, om die risiko van siekteherhaling in borskanker oorlewendes as gevolg van mislukking van kankerbehandeling wat nie in ooreenstemming is met die pasiënt se genotipe nie, te voorkom. Die tweede is die afbakening van sleutel paaie waardeur gene wat betrek word by borskanker en verwante ko-morbiditeite, voedings- en medikasie behandelingsteikens oor diagnostiese grense heen kan voorsien. Die derde is die ontdekking van genetiese veranderinge onderliggend aan familiële borskanker wat nie verklaar kan word deur mutasies in die twee belangrikste kanker-onderdrukker gene, BRCA1 en BRCA2 nie.

Die studiepopulasie het bestaan uit 164 borskankerpasiënte (60 Kleurling/Gemengde Afkoms en 104 Koukasiers), van wie 88 geselekteer is uit 'n totaal van 813 individue wat ingeligte toestemming verleen het vir insluiting van hulle data in ‘n genomiese databasishulpbron, wat ontwikkel is op die interfase van die laboratorium en roetine kliniese praktyk. DNA van ‘n verdere 101 kankervrye individue bokant die ouderdom van 65 jaar was beskikbaar vir kliniese validasie van potensieel kanker-verwante variante in ‘n uitgebreide vroulike kontrole groep.

In die eerste fase van die studie is reël-tyd polimerase ketting reaksie (PKR) TaqMan© tegnologie gebruik om die potensiële waarde te bepaal om farmakogenetiese toetsing (CYP2D6 alleel 4) by te voeg tot standaard immunohistochemie (IHC)-gebasseerde borstumour subtipering wat gekomplementeer is deur BRCA mutasie sifting en/of mikroarray-geen profilering in toepaslike pasiënte. In fase twee van die studie is algemene genetiese risikofaktore vir kardiovaskulêre siekte (KVS) aangetoon om betekenisvol geassosieerd te wees met vroeër ouderdom (10 jaar gemiddeld) van borskanker aanvang/diagnose (APOE E4 alleel, p=0.003; 95% CI: 4-15) en liggaamsmassa indeks (BMI) (MTHFR 1298 A>C; p=0.01; 95% CI: 3-14) in pasiënte wat gestratifiseer is volgens estrogeen reseptor (ER) status, na aanpassing vir potensiële strengelveranderlikes. Ouderdom by diagnose/aanvang van borskanker was betekenisvol laer in ER-negatiewe as ER-positiewe pasiënte na statistiese aanpassing vir etnisiteit (p=0.022), terwyl BMI beduidend hoër was in ER-positiewe as ER-negatiewe borskanker pasiënte na aanpassing vir ouderdom, etnisiteit en familiegeskiedenis van kanker (p=0.035). Hierdie bevindinge het bygedra tot die ontwikkeling van ‘n eksoom pre-siftings algoritme (EPA) wat gebruik is in deel 3 van hierdie studie om drie geneties ongekarakteriseerde borskanker pasiente vir WES te selekteer. Die resultate is vergelyk met die bevindings in drie kankervrye kontroles uit dieselfde populasiegroep. WES, gevolg deur variant

roeping met beide die standaard mensgenoom verwysingsvolgorde (hg19) en ‘n etnies verenigbare major alleel verwysings genoom (MARS) het meer as 20% teenstrydigheid getoon in die aantal geen variante wat geïdentifiseer is in dieselfde DNS monsters. Nadat ‘n groot aantal vals-positiewe resultate uitgesluit is wat deur minor allele in hg19 veroorsaak is, is twee raar missens mutasies (<1%) geïdentiseer in ‘n familie met ER-positiewe borskanker: RAD50 R385C en MUC1 Q67E. Drie verskillende bioinformatika programme is gebruik om funksionaliteit te voorspel en albei mutasies is bevestig deur Sanger volgorderbepaling en/of RT-PKR. Nie een van die twee missens mutasies in die RAD50 of MUC1 gene is gevind in die eksome van ‘n onverwante borskanker pasiënt met trippel-negatiewe borskanker of in die eksome van die drie kontrole indiwidue van dieselfde populasiegroep nie.

Hierdie studie het gelei tot die ontwikkeling van 'n raamwerk vir WES wat n verbetering is op die beperkings van enkelgeen BRCA-toetsing van Suid-Afrikaanse borskankerpasiënte. Ons bevindinge ondersteun vorige WES-resultate wat aangedui het dat die meeste geneties-ongekarakteriseerde familiere borskankers waarskynlik veroorsaak word deur 'n kombinasie van lae tot matige penetrasie-mutasies wat tot uitdrukking kom in 'n hoë-risiko omgewing soos gereflekteer word deur hoë BMI. WES maak dit moontlik om genetiese risikofaktore te identifiseer wat relevant is tot beide kankerontwikkeling en geteikende terapeutiese intervensie met dieselfde genetiese toets.

Acknowledgements

I would like to take this opportunity to acknowledge and express my sincere gratitude toward the following individuals and institutions without whom the success of this study would not have been possible:

Firstly, to my supervisor, **Professor Maritha J. Kotze** who has played an instrumental role in the write-up of this dissertation and my overall development as a young scientist - thank you for your expert supervision and for granting me the opportunity to become part of the research community and gain knowledge and exposure in the field through attending both local and international conferences. Thank you for seeing my potential, acknowledging my strengths and allowing me to grow in the areas of my interest. Your supreme knowledge and ideas as well as continued support, guidance and patience, have ensured the optimal learning and uplifting experience. Your passion, dedication, out-of-the-box thinking and willingness to challenge the seemingly impossible, is truly inspiring.

Secondly, my **co-supervisors** Dr Rika Pienaar, Professors Juanita Bezuidenhout and Susan J van Rensburg for their support and expertise related to clinical, pathology and biochemistry information included in this study. Dr Rika Pienaar is especially thanked for referring some of the patients included in this investigation.

A special thanks to my **family** (dad, mom, sister, brother and grandparents) for enduring neglect, hours of silence, as well as impatience and sometimes moodiness during stressful periods. I am grateful for their encouragement and constant moral support, and humbled by their love, acceptance of, and unwavering belief in me. To m y **friends** and extended family, your concern and words of encouragement meant a lot to me during the times when I needed it most; thank you.

Dr Hilmar K. Luckhoff is acknowledged and thanked for his availability and readiness to assist in the editing of my manuscript and his contribution toward the statistical analyses performed in this study. His invaluable input and the mere knowledge that he's in my corner, is deeply appreciated.

The Central Analytical Facility and specifically the manager of the DNA Sequencing Unit, Carel van Heerden, is thanked for the Sanger and whole exome sequencing performed on the samples of patients included in this study. The Pathology Research Facility (PRF) is thanked for

the use of the standard operating protocols (SOPs) related to the assays included in the CVD multi-gene test performed as part of the chronic disease/wellness screen used to develop the Gknowmix database used in this study to validate the variants identified with the use of whole exome sequencing. In particular, **Leslie Fisher, Kelebogile Moremi and Dr Armand Peeters** are acknowledged and thanked for their specific contributions in relation to primer design, conventional and Real-time PCR experiments performed in the PRF laboratory. Your patience and time invested in this context is highly appreciated.

Mr Dieter Geiger is greatly thanked for his support and belief in me throughout my postgraduate studies. His critical thinking taught me that, more important than finding the correct answers, is asking the right questions. His wisdom and high-spirited nature truly enriched my journey.

Professor Lize van der Merwe of the Biostatistics Unit of the MRC is tremendously thanked for performing the statistical analysis of the study, occasionally at her own inconvenience.

The breast cancer **patients** and control individuals are thanked for their consent to participation in this project.

The Strategic Health Innovation Partnerships (SHIP) Unit of the South African Medical Research Council (MRC), Technology Innovation Agency (TIA), Cancer Association of South Africa (CANSA), Winetech and Technology for Human Resources and Industry Program (THRIP) are acknowledged and thanked for financial assistance.

The Department of Pathology, Division of Anatomical Pathology at Stellenbosch University is thanked for supplying the infrastructure required to complete this project.

The Creator for giving me the strength, knowledge, perseverance and the enthusiasm to learn.

Dedication

I would like to dedicate this thesis to my parents, who taught me that;

“When you aim for the moon and miss, you’ll still be amongst the stars”.

Table of Contents

Abbreviations and Symbols

5’	5-prime
3’	3-prime
α	alpha
β	beta
©	copyright sign
°C	degrees Celsius
=	equal to
>	larger than
µg/L	microgram per litre
µl	micro litre
-	minus
%	percentage
+	plus
±	plus-minus
®	registered trademark
<	smaller than
A	adenine
A (Ala)	alanine
AAP(s)	atypical antipsychotic(s)
AD	Alzheimer’s disease
ATM	
ATP	adenosine 5’-triphosphate
bp	base pair
BC	breast cancer
BFAST	Blat-like Fast Accurate Search Tool
BLAST	basic local alignment search tool
BRCA1/2	breast cancer gene 1/2
BWA	Burrows-Wheeler aligner
BWT	Burrows-Wheeler transform
C	cytosine
C (Cys)	cysteine

Ca	cancer
CNV	Copy Number Variation
CYP2D6	cytochrome P450 2D6
CVD	Cardiovascular disease
D (Asp)	aspartic acid
dATP	2'deoxy-adenosine-5'triphosphate
dCTP	2'deoxy-cytosine-5'triphosphate
ddATP	2',3'-dideoxy-adenosine-5'triphosphate
ddCTP	2',3'-dideoxy-cytosine-5'triphosphate
ddGTP	2',3'-dideoxy-guanosine-5'triphosphate
ddH_2O	double distilled water
ddTTP	2',3'-dideoxy-thymidine-5'triphosphate
dGTP	2'-deoxy-guanosine-5'-triphosphate
dH_2O	distilled water
DNA	deoxyribonucleic acid
dsDNA	double stranded DNA
dbSNP	The Single Nucleotide Polymorphism Database
dTTP	2'-deoxy-thymidine-5'-triphosphate
EDTA	ethylenediaminetetraacetic acid
ER	estrogen receptor
EtBr	ethidium bromide
FRET	fluorescence resonance energy transfer
g	gram
G (Gly)	glycine
G	guanine
GATK	genome analysis tool kit
GSTT1	Glutathione S-transferase
GWAS	genome wide association study
H (His)	histidine
H_2O	water

H_3BO_3	boric acid
I (Ile)	isoleucine
IVS	intervening sequence
L (Leu)	leucine
LiPA	reverse hybridization line-probe assay
M (Met)	methionine
M	molar
mg	milligram
$MgCl_2$	magnesium chloride
ml	millilitre
mM	milli-molar
MAF	minor allele frequency
MRI	magnetic resonance imaging
mRNA	messenger ribonucleic acid
MTHFR	methylenetetrahydrofolate reductase
MTR	methionine synthase
MTRR	methionine synthase reductase
N (Asn)	asparagine
NaCl	sodium chloride
ng	nanogram
ng/µl	nanogram per micro litre
NCBI	National Centre for Biotechnology Innovation
NGS	Next Generation Sequencing
nsSNP	nonsynonymous Single Nucleotide Polymorphism
OMIM	Online Mendelian Inheritance in Man
OCP	Oral contraceptive pill
p	short arm of chromosome
P (Pro)	proline
PCR	polymerase chain reaction
PD	Parkinson’s disease

PM	poor metabolizer
pmol	picomole
PR	progesterone receptor
q	long arm of chromosome
Q (Glu)	glutamine
QC	quality control
Q score	Phred score
R (Arg)	arginine
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RT PCR	real-time polymerase chain reaction
rxn	reaction
S (Ser)	serine
SIFT	Sorting Intolerant From Tolerant
SNP(s)	single nucleotide polymorphism(s)
SNV	single nucleotide variant
sSNP	synonymous Single Nucleotide Polymorphism
SSRIs	selective serotonin reuptake inhibitors
T (Thr)	threonine
T	thymine
T_A	annealing temperature
TAM	tamoxifen
Taq	*Thermus aquaticus* polymerase enzyme
TBE	Tris-Borate-EDTA buffer
TCA	tricyclic antidepressants
TE	Tris-EDTA buffer
T_M	melting temperature
TM	trademark

u	units
uTR	untranslated region
UV	ultraviolet
V (Val)	valine
V	volts
v/v	volume per volume
VUS	variants of unknown clinical significance
w/v	weight per volume
WES	whole exome sequencing
WGS	whole genome sequencing
x	times
x g	times gravity
Y (Tyr)	tyrosine

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

Molecular pathology plays an increasingly important role in cancer prevention, diagnosis and treatment. The significant impact of breast cancer on the health economy sparked extensive research in South Africa (Grant et al. 2013; Grant et al. 2014) and elsewhere into the molecular pathways underlying distinct tumour subtypes that differ in their response to treatment. In South Africa breast cancer is the most common neoplasm among Asian and Caucasian women, and the second most common cancer among women of African descent and the Coloured population of Mixed Ancestry (Vorobiof et al. 2001). Inter-ethnic variability in breast cancer risk is explained by a combination of genetic and environmental risk factors that contribute to cancer development and tumour gene expression.

Over the last decade a new era of personalized medicine was introduced with the promise that pharmacogenomics will deliver the right treatment at the right time in patients with cancer and associated diseases (Evans and Relling 2004). As is evidenced by the exponential growth in applied pharmacogenetics literature over the past 10 years (Figure 1.1), pharmacogenetics has become one of the leading and potentially most actionable areas of the personalized medicine paradigm (Frueh et al. 2008; Lesko and Zineh 2010; Zineh and Pacanowsi 2011).

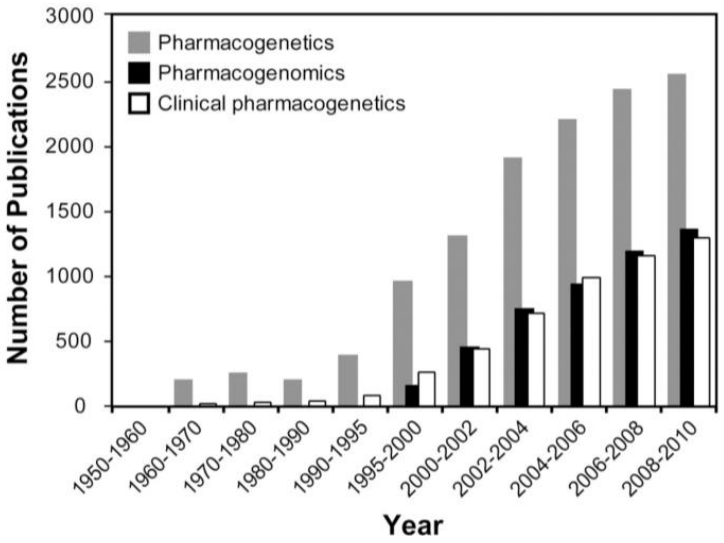


Figure 1.1. Number of Pubmed citations (http://www.ncbi.nlm.nih.gov/pubmed) by date using the keywords “clinical pharmacogenetics”, “pharmacogenomics” or “pharmacogenetics”. Reproduced with permission from Scott (2011).

The extent to which breast cancer gene 1 and breast cancer gene 2 (BRCA1 and BRCA2) mutations, non-deterministic genetic risk modifiers and relevant lifestyle factors influence individual predisposition towards the development of breast cancer in a subtype-specific

manner, represents an important avenue for ongoing research endeavors. Most studies to date have failed to demonstrate a significant overall survival difference between sporadic and familial breast cancer caused by mutations in the BRCA1 and BRCA2 genes (Bordeleau et al. 2010). This finding confirms the importance of genetic risk modifiers and lifestyle factors as determinants of adverse clinical outcome across breast cancer subtypes, including BRCA-positive patients (King et al. 2003). Newman et al. (2008) furthermore provided compelling evidence that treatment failure due to impaired drug metabolism caused by reduced Cytochrome P450 family 2 subfamily D polypeptide (CYP2D6) enzyme function, may increase the risk of recurrence in patients with BRCA mutations or other forms of familial breast cancer.

The above findings are of particular relevance in South Africa due to a founder effect found to be responsible for a relatively high frequency of specific BRCA1/2 mutations in high-risk families with multiple affected members (Reeves et al. 2004; Agenbag 2005; van der Merwe et al. 2012b; van der Merwe and van Rensburg 2009). This prompted an investigation into the potential significance of combining diagnostic BRCA1/2 and predictive CYP2D6 pharmacogenetic testing in South African breast cancer patients, as reported in Phase 1 of this study (van der Merwe et al. 2012a). CYP2D6 metabolises more than 25% of all commonly prescribed drugs and may therefore provide a treatment target across diagnostic boundaries. This is of particular relevance in patients with breast cancer and associated co-morbidities such as depression and cardiovascular disease (CVD) due to the high frequency of therapeutic failure reported with use of antidepressants (as substrates or inhibitors of CYP2D6 enzymatic activity), cholesterol-lowering statins, beta-blockers, and tamoxifen (Frudakis et al. 2007; Newman et al. 2008; Nagele et al. 2011; Suppli et al. 2011). Tamoxifen has been associated with the development of deep vein thrombosis (DVT) and various cognitive effects (Paganini-Hill et al. 2000; Eberling et al. 2004; Decensi et al. 2005). Although the evidence regarding association of tamoxifen with depression is inconsistent, mood changes are frequently reported with the use of this hormonal anti-cancer treatment (Thompson et al. 1999; Patten and Barbui 2004).

Previous studies have shown that the risk for developing CVD may exceed the risk of breast cancer recurrence in survivors (Patten et al. 2012). Obesity which represents a major metabolic risk factor for CVD is reciprocally associated with depression (Faith et al. 2002; de Wit et al. 2009) and increases the risk of various cancers, including postmenopausal breast cancer, colorectal and prostate cancer (Louie et al. 2013). Li et al. (2009) found a 40-50% elevation in risk associated with obesity, which is consistent with the results from women enrolled in the National Surgical Adjuvant Breast and Bowel B-14 randomized trial of adjuvant tamoxifen performed in patients with node-negative estrogen receptor (ER)-positive tumours (Dignam et

al. 2003). Breast cancer survivors in the highest quartile for body mass index (BMI) had at least 50% increased risk of a second primary breast cancer compared to women in the lowest BMI quartile (Trentham-Dietz et al. 2007). The negative effect of obesity and lack of exercise also applies to familial breast cancer caused by mutations in the BRCA1 and 2 genes (King et al. 2003). Swisher et al. (2015) confirmed the benefits of exercise and a healthy diet to improve the quality of life in survivors of triple-negative breast cancer, the tumour subtype most frequently associated with mutations in the BRCA1 gene.

Obesity represents an important target for lifestyle intervention and application of pharmacogenetics related to the use of anti-inflammatory drugs has been shown to reduce the risk of cancer (Ford et al. 2015). It has been estimated that 15-20% of all cancer deaths may be attributable to obesity (Calle et al. 2003; Chang et al. 2013). Since tumour growth is regulated by interactions between cancer cells and their tissue microenvironment, dysfunctional fat tissue and altered signaling pathways were identified as important mechanisms whereby obesity could lead to cancer development. Duval and Tweedie (2000) showed that post-menopausal obese women possessed a three-fold increase in the risk of breast cancer compared to their non-obese control counterparts. Furthermore, Laumann et al. (2008) has shown that a higher BMI is associated with worse response to neoadjuvant chemotherapy and worse overall survival.

Genes involved in the folate-homocysteine pathway have been identified as important therapeutic targets in relation to both gene-diet and gene-diet interaction, due to the known deleterious effects of obesity, excessive alcohol intake and smoking on DNA methylation which is implicated in the development of breast cancer and associated comorbidities. Decreased genomic DNA methylation may occur as a consequence of reduced enzyme activity caused by low folate status and polymorphic variation in at least 30 genes involved in the folate metabolism pathway (Friso et al. 2002; Wettergren et al. 2010). These include the most extensively studied methylenetetrahydrofolate reductase (MTHFR), methionine synthase (MTR) and methionine synthase reductase (MTRR) genes. The enzymes encoded by the aforementioned genes are important in the one-carbon transfer processes known to be essential for optimal DNA synthesis, methylation and repair. The methylation pathway therefore presents an important target for possible prevention of cumulative effects that could impact on the development of comorbidities or disease relapse in breast cancer patients (Peerbooms et al. 2011; Cohen et al. 2012). We have previously studied the clinical relevance of the MTHFR 677 C>T mutation in South African individuals participating in a chronic disease screening program and could confirm the association between low dietary folate intake, obesity and depression (Delport et al. 2014).

The development of new ways to use current genomic knowledge for the benefit of patients remains a challenge in the field of breast cancer. Albeit significant advances over recent years, screening for causative mutations underlying familial breast cancer not attributed to defects in the BRCA1 and BRCA2 genes revealed many variants of uncertain clinical significance (VUS). This presents a dilemma not only in treatment decision-making, but also in relation to genetic counselling of patients and their at-risk family members. Mutations in the BRCA1 and BRCA2 genes account for less than 50% of familial breast cancer (Wooster et al. 2003), which suggest that other genes, individually or in combination, confer the residual risk seen in heritable cancers. Our understanding of the germline contribution to gene expression levels in tumours have improved significantly with the use of next generation sequencing (NGS). This technology enabled the identification of rare gene variants with high impact as well as low-moderate-penetrance mutations, which could collectively account for disease risk and treatment failure in genetically uncharacterized patients. Sequencing of the entire BRCA1 and 2 genes in South African patients is usually only performed after exclusion of founder mutations causing protein truncating mutations (Reeves et al. 2004; Agenbag 2005; van der Merwe et al. 2012b; van der Merwe and van Rensburg 2009). Although detection of causative missense mutations in the BRCA1/2 genes are uncommon, BRCA1 R1699Q (c.5096G>A, p.Arg1699Gln) classified as an intermediate-risk allele (Spurdle et al. 2012) has recently been reported in a South African family with breast cancer (unpublished data). It is therefore important not to disregard missense mutations in the BRCA1/2 or other cancer-related genes without further studies to determine possible clinical relevance.

The vast amounts of information that have been generated with the advent of NGS currently exceed the rate at which large amounts of data are validated in a diagnostic setting. Given the slow pace of validation studies, translation of genomic discoveries into practical benefits may be one of the most difficult challenges to meet at present. A genomics database resource for determination of the clinical relevance of new genomic information is therefore required for on-going assessment and monitoring of treatment outcomes.

Early detection of relatively frequent (>1%) functional polymorphisms with low penetrance as putative actionable treatment targets may facilitate prevention of cumulative risk across the disease spectrum. Single nucleotide polymorphisms (SNPs) in genes encoding drug-metabolizing enzymes and those dependent on vitamin co-factors for optimal function may explain inter-individual variation in treatment response, drug toxicity and/or recurrence risk in both familial and sporadic breast cancer. Recognition of the significant impact of multi-functional SNPs found to be of clinical relevance in cancer and related comorbidities support the

development of a pharmacogenomics assay that could serve as a prescreen step for whole exome sequencing (WES) or targeted NGS.

1.1. Research focus

The current investigation is based on the hypothesis that integration of genetic information with clinical and environmental factors (e.g. medication, lifestyle) may allow the development of a novel pathology-supported pharmacogenomics assay for improved clinical management of breast cancer and associated comorbidities. Although genetic as well as modifiable environmental risk factors including obesity, alcohol consumption and smoking were shown to have a significant effect on the recurrence rate of breast cancer (Li et al. 2009), their role in ER-positive versus ER-negative breast cancer patients remains to be elucidated in the South African context.

Although the focus of personalized genomic medicine is on disease management, the potential for cancer prevention calls for a wider clinical application in primary healthcare. Technological advances derived from whole genome sequencing (WGS) and genome wide association studies (GWAS) provides the capacity to integrate clinical and genotypic information for assessment of generalized health risks and promotion of wellness (Patel et al. 2013). Current health promotion strategies incorporate lifestyle factors and clinical indicators such as BMI into screening programs, but are largely insufficient to evaluate the role of medication side effects and drug failure in the development of comorbidities and recurrence risk among cancer survivors. The potential value of combining all these factors with genomic information using a comprehensive pathology-supported genetic testing (PSGT) strategy developed in South Africa over recent years (Kotze et al. 2015) has not previously been explored in breast cancer patients.

The assessment of non-deterministic risk modifiers incorporated into a multi-gene assay that also takes BRCA mutation status into account may be ideally suited to determine eligibility for WES using the PSGT approach. Towards this goal, three challenges were identified and addressed in this study. Firstly, the need to combine diagnostic BRCA mutation detection with pharmacogenetic testing to reduce recurrence risk in familial breast cancer due to inappropriate treatment. Secondly, the delineation of pathways through which genes implicated in both breast cancer and related co-morbidities may be targeted for individualized nutritional and drug treatment across diagnostic boundaries. Thirdly, the discovery of genetic alterations underlying familial breast cancer not attributed to mutations in BRCA1 and BRCA2 and determination of clinical relevance. While NGS provides a way to cover all three areas of risk in one test, new

ways to interpret the vast amount of new genomic knowledge derived from WES is required for clinical application to the benefit of each individual patient.

Current research initiatives acknowledge the need for defining selection criteria to identify genetic subgroups of patients where the potential benefits of genetic testing would outweigh the risk of unnecessary testing (Fostira et al. 2012). Identification of disease subtypes with a genetic component allows for classification of patients into different treatment groups based on disease mechanism and not only symptoms (Phipps et al. 2011; Phipps et al. 2012). Development of a genetic testing strategy whereby pathology and genetic tests are assessed together in a multiplex assay may overcome the limitations of individual health disciplines, especially if combined with bioinformatics tools to apply WES in genetically uncharacterized patients with familial breast cancer. To bridge the gap between the generation of massive amounts of genomic information towards achievement of a high degree of concordance between genotype and phenotype, the generation of a database resource for validation of new genomic applications was a priority.

The aim of this study was to provide a method to facilitate the prevention of cumulative risk for comorbidities such as vascular disease and depression in breast cancer patients through the development of a comprehensive PSGT service informed by molecular tumour subtyping. Results from histopathological assessment and genetic testing were assessed as part of a multidisciplinary approach to chronic disease screening used to identify patients eligible for extended genomic analysis including microarray-based gene expression profiling (Grant et al. 2013; Grant et al. 2014) and NGS (this study). The etiology, epidemiology, pathology and treatment of breast cancer, as well as the different categories of risk in relation to testing (e.g. diagnostic, predictive/pharmacogenetics), exome sequencing, and counselling aspects related to the genomics era are reviewed in chapter 2. The aims of the study are motivated by the identification of gaps in current knowledge (chapter 3) related to three challenges addressed in three phases of the current investigation:

In **Phase I** of the study, we determined the appropriateness of adding CYP2D6 pharmacogenetic testing to the breast cancer genetic screening options already available in South Africa, which includes BRCA mutation screening and transcriptional profiling to assess ER, PR and HER2 status.

In **Phase II**, high BMI and other CVD risk factors were assessed in breast cancer patients, stratified according to ER status, to facilitate the development of an exome pre-screening algorithm (EPA).

In **Phase III**, the newly developed EPA was used to select three breast cancer patients (from two unrelated families) for whole exome sequencing (WES) in comparison with three cancer-free controls. Discordance reported between different variant calling pipelines prompted comparison between the standard human reference genome (hg19) and a synthetic major allele reference sequence (MARS).

A description of the study population and methodology employed is documented in chapter 4 while the findings obtained in relation to the Phase 1-3 analyses are presented separately in the results and discussion sections (Chapter 5). This is followed by a general conclusion based on the new knowledge and the database resource generated as a result of this investigation (Chapter 6). The references used (Chapter 7) is followed by additional results pertaining to each of the study phases and concluded with a list of congress presentations and publications in an appendix chapter (8).

CHAPTER 2

LITERATURE REVIEW

2.1. Breast cancer evolution

Breast carcinoma is a form of genetic disease that originates from breast tissue. Records and illustrations of breast cancer date back to ancient times, approximately 3000- 2500 B.C., with The Edwin Smith Surgical Papyrus providing accurate accounts of the disease (Breasted 1984). In 400 B.C., Hippocrates described the progressive stages of breast cancer and developed the humoral theory which ascribed the imbalance of liquids in the body as a cause of disease, representing early hypotheses on the etiology of cancer. He believed that an imbalance of the 4 bodily fluids (blood, phlegm, yellow bile and black bile), particularly an excess of black bile underlie carcinogenesis (Homer 1966). In the 1st century A.D., Leonides of Alexandria described his approach of incision and cautery (Ariel 1987), which involved leaving a wide margin of excision and only removing relatively small tumours, introducing the principles of modern-day surgical practice (Lewison 1953). Galen in 200 A.D. on the other hand contended that breast cancer was a systemic disease and attributed it to the accumulation of black bile in the blood (De Moulin 1983). The aforementioned physicians proposed that menopause was somehow linked to cancer, indirectly associating the disease with old age.

2.2. Etiology of breast cancer

Carcinogenesis is a multi-step process, the etiology of which includes both environmental factors such as tobacco smoke, chemicals, radiation, and infectious agents and internal factors including hormones, immune conditions, and genetic variations. These factors may act in concert to initiate or promote the development of neoplasms resulting from dysfunction of regulation of cell growth due to the accumulation of genetic alterations (Garcia et al. 2007; Stratton et al. 2009). Malignant transformation involves a progressive process starting with atypical hyperplasia followed by carcinoma in situ and eventually evolution from pre-invasive to invasive breast cancer. Benign lesions differ from malignant tumours in terms of cellular growth rate (relatively slow growth) and morphology (tumour cells resemble in appearance that of normal cells) and are not life-threatening. Pre-cancerous lesions including atypical hyperplasia and carcinoma in situ do not invade the surrounding fatty breast tissue or spread to lymph nodes or distant sites. Malignancy on the other hand is characterized by dissemination of tumour cells beyond the original tumour to other parts of the body (American Cancer Society 2009). The spread of these unregulated cells (metastasis) may be fatal if growth is not controlled or ceased.

2.3. Histopathological classification of breast cancer

Breast cancer includes a collection of different diseases or subtypes characterized by different biological and pathological features, clinical presentation, response to treatment, clinical behavior, and outcome (Rivenbark et al. 2013; Dieci et al. 2014). Heterogeneity is observed on different levels ranging from the classic histopathological characterization to the more modern molecular classification. Using TNM (tumour size, nodal spread, metastases) staging, the disease can be broadly categorized into in situ carcinoma and invasive (infiltrating) carcinoma. Breast carcinoma in situ is further sub-classified as either ductal or lobular whereby growth patterns and cytological features form the basis to distinguish between the two types (Malhotra et al. 2010). Invasive ductal carcinoma (IDC) that develops in the milk ducts of the breast is the most common type of breast cancer, comprising approximately 65-85% of all breast cancers (Figure 2.1).

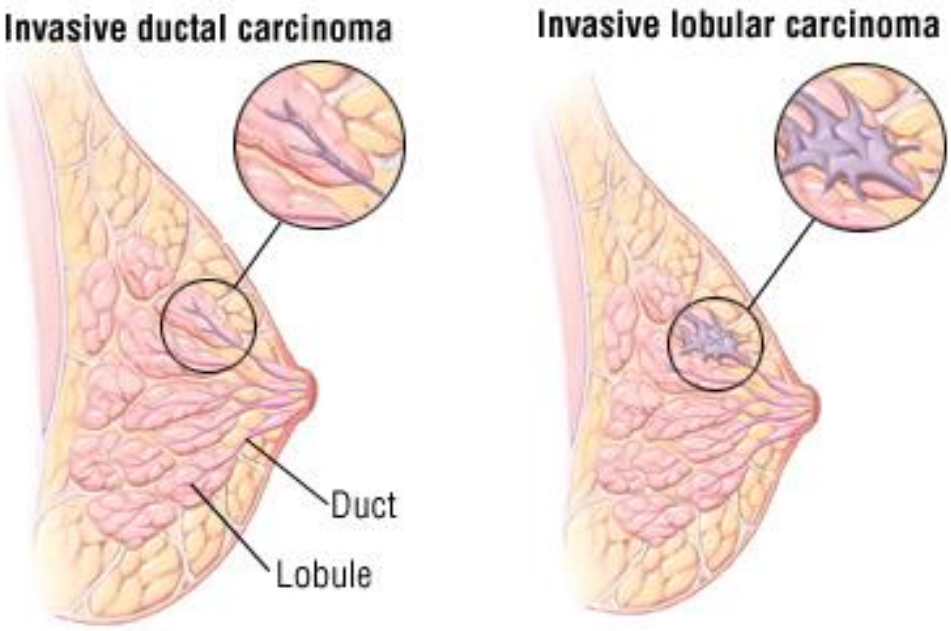


Figure 2.1. Figure depicting invasive ductal and invasive lobular carcinoma (http://www.drugs.com/health-guide/breast-cancer.html).

Less commonly diagnosed is invasive lobular breast carcinoma (ILBC) that develops in the breast milk-producing lobules or glands and comprise about 10-15% of all cases of advanced breast cancer (Leong and Zhuang 2011; Onitilo et al. 2009). Other malignancies of the breast include nipple cancer and relatively rare undifferentiated neoplasms (American Cancer Society 2015). Ductal and lobular carcinoma in situ (DCIS and LCIS) are considered precursors or predisposing determinants of risk for subsequent invasive disease. In 2003, the overall 5-year

survival rate for both IDC and ILC was approximately 85% (Arpino et al. 2004). DCIS in itself is harmless, although if untreated approximately 60% of these low-grade lesions will become invasive over the course of 40 years follow-up (Evans et al. 2004). Table 2.1 provides a summary of the various types of breast cancer, their anatomical sites of origin and prevalence.

Table 2.1. Classification and prevalence of the different types of breast cancers.

Breast cancer type	Anatomical site	Prevalence
Ductal Carcinoma	Milk ducts	73%
Invasive ductal carcinoma (IDC)	Cancer originates in milk ducts but invades breast tissue	65-85%
Ductal carcinoma in situ (DCIS)	Grows only on the lining of milk ducts	15%
Lobular Carcinoma	Lobes or lobules of the breast	12%
Invasive lobular carcinoma (ILC)	Cancer originates in lobules but invades breast tissue	10-15%
Lobular carcinoma in situ (LCI)	Grows only in the lobes and terminal ducts of the breast	Unknown
Other (nipple, mucinous, medullary, tubular, adenoid cystic, angiosarcoma, phyllodes, papillary, metaplastic, lymphoma)	Various sites including connective tissue and lymph nodes	~10%

2.4. Disease Prevalence

Cancer is the leading cause of death in first world countries and the second leading cause of death in developing countries (World Health Organization 2008). Breast cancer is the most common non-cutaneous neoplasm in females, with more than 1 million women diagnosed worldwide and about 400 000 women killed each year (Kamangar et al. 2006). In 2012, 1.7 million women received a diagnosis of breast cancer. Since 2008 mortality has increased by 14% while estimates of breast cancer incidence have increased by more than 20% worldwide (International Agency for Research on Cancer 2013). In South Africa more than 6 000 women are diagnosed with breast cancer annually and approximately 40% of them are predicted to die from the disease (http://www.cansa.org.za). Breast cancer is second only to lung cancer as the most common cause of cancer mortality (Ferlay et al. 2013). The disease is ~100 times more

common in women than it is in men, however, males tend to have poorer outcomes due to delay in diagnosis. Although incidence rates have risen approximately 30% in westernized countries over the preceding 30 years, mortality rate reductions have been attributed to earlier detection of the disease through mammography and screening programs as well as improved treatment strategies. In contrast, incidence rates have been rising in many developing countries including Asian and African countries (Ngoma 2006).

According to the American Cancer Society (2015), an estimated 231,840 new cases of invasive breast cancer are expected to be diagnosed among women in the United States during 2015, with approximately 2,350 men affected. Among Caucasian women, the breast cancer incidence rate decreased by nearly 7% from 2002 to 2003. This decline has been attributed to reductions in the use of hormone replacement therapy (HRT), subsequent to reports that stated that use of combined estrogen plus progestin therapy was associated with an increased risk of breast cancer (Gann and Morrow 2003), venous thrombosis (Cushman et al. 2004), coronary heart disease (Manson et al. 2003), dementia and mild cognitive impairment (Shumaker et al. 2003). In contrast, HRT use has been associated with reduction of the risk of osteoporosis (Stevenson 2006; Marjoribanks et al. 2012) and more controversially an increased risk of CVD (Hickey et al. 2012; Schierbeck et al. 2012). Incidence rates which were stable in Caucasian women, marginally increased in black women (0.3% per year) between 2007 and 2011.

Due to these alarming figures and the heterogeneity of the South African population, a need exists for the identification of subgroups of the population that may be at increased risk of breast cancer development or recurrence due to genetic susceptibility, environmental exposures or treatment failure.

2.5. Risk factors

The primary breast cancer risk factors that have been identified are a genetic predisposition, age, gender, lack of childbearing or breastfeeding, and higher hormone levels (Giordano et al. 2004; Collaborative Group on Hormonal Factors in Breast Cancer 2002, Yager and Davidson 2006). Reproductive risk factors include a long menstrual history, use of oral contraceptive pills, and older age at first child birth (>30 years) (Warwick et al. 2003). Together, the increased breast cancer risk conferred by the aforementioned factors may reflect cumulative estrogen effects on breast tissue.

Substantial clinical and epidemiological evidence have linked cumulative and sustained exposure to estrogens with increased risk of developing breast cancer. The most established risk factors for breast cancer are thought to influence risk through increasing concentrations of

endogenous estrogens (Pike et al. 1998). Moreover, the most widely acknowledged mechanism of estrogen carcinogenicity is binding to its specific nuclear receptor alpha, thereby exerting a potent stimulus on breast cells to proliferate. Studies that isolated the estrogen receptor (ER) protein to correlate ovarian function with estrogen production, mutually led to the identification of a strong association between estrogen dose/length of exposure and increased breast cancer risk (Cauley et al. 1999; Kabuto et al. 2000). Estrogens are considered to play a major role in promoting the proliferation of both the normal and the neoplastic breast epithelium. Several mechanisms are postulated to be involved in their carcinogenic effects: stimulation of cellular proliferation through their receptor-mediated hormonal activity, direct genotoxic effects by increasing mutation rates (oxidative DNA damage that cause DNA breaks) and induction of aneuploidy (Russo et al. 2006; Giam and Rancati 2015).

Moreover, a family history of breast cancer and certain breast stages giving rise to abnormal cells confers increased risk for breast cancer. A woman's risk is significantly increased if a 1st or 2nd degree family member has had early-onset breast cancer (<40 years) (American Cancer Society 2015). Ethnicity also plays a role in the development of breast cancer as risk differs across ethnic groups (Levy-Lahad et al. 1997; Reeves et al. 2004; Agenbag 2005, van der Merwe and van Rensburg 2009, van der Merwe et al. 2012b).

Non-modifiable factors associated with increased breast cancer risk include high breast tissue density (the ratio of glandular tissue to fatty tissue, measured on a mammogram) (Whitehead et al. 1985), high bone mineral density (Qu et al. 2013), type 2 diabetes (Boyle et al. 2012), high-dose radiation to the chest as a result of medical procedures and benign breast conditions such as biopsy-confirmed atypical hyperplasia, DCIS and LCIS (American Cancer Society 2015).

In addition, the use of post-menopausal HRT (Ross et al. 2000), lack of physical activity, obesity, smoking and consumption of one or more alcoholic beverages per day have been associated with increased breast cancer risk (Room et al. 2005; American Cancer Society 2015). In 2010, a study conducted on 72 000 women by the National Cancer Institute (NCI) found that compared to women who maintained their weight, 20-year old women who progressively gained weight as they aged doubled their risk of developing breast cancer after menopause. In addition, overweight (BMI ≥ 25) or obesity (BMI ≥ 30) adversely affects survival in postmenopausal patients (Chlebowski et al. 2011). Overweight and obesity are responsible for 6.9% of cancers, which is nearly double the proportion caused by infectious agents estimated as 3.7%. Li et al. (2009) demonstrated a 7.2-fold elevated risk of contralateral breast cancer in women who smoked and consumed 7 or more alcoholic beverages per week. Between 4-5% of breast cancer cases are primarily caused by alcohol consumption

(Longnecker 1994; Parkin et al. 2011). Smoking and alcohol consumption were not associated with increased risk of triple negative breast cancer (TNBC), but were modestly associated with increased risk of ER-positive breast cancer (Kabat et al. 2011). A pooled analysis of >300 000 women and 4 335 cases suggested that consumption of 2-5 alcoholic beverages per day increases the risk of breast cancer by roughly 40% (Smith-Warner et al. 1998). The underlying mechanisms through which alcohol is thought to increase the risk for breast cancer include an influence on circulating estrogen levels (Purohit 1998), enhanced permeability of chemical carcinogens, immune function, metabolism of alcohol to acetaldehyde (a known carcinogen) (Feron et al. 1991) and decreased absorption of essential nutrients such as dietary folate (Thomas 1995). Diet is of particular relevance to this study which focuses on gene-diet interaction (nutrigenetics) that may modify the risks associated with breast cancer. The Mediterranean diet has been linked to a reduction in the risk of breast cancer (Toledo et al. 2015).

Breast feeding, reduced alcohol consumption, moderate or vigorous physical activity and maintaining a healthy body weight have been associated with lower risk of breast cancer (American Cancer Society 2015).

2.6. Breast cancer screening and prognosis

Early detection remains the primary defense available to patients in preventing the development of breast cancer. Despite advances in imaging technology, discrepancies regarding recommended schedules have complicated the process of screening. Careful consideration therefore needs to be given to the benefits and harms of the screening intervention, along with the costs involved.

Breast cancer screening refers to the evaluation of healthy women (showing no signs or symptoms of the disease) for breast cancer in an attempt to achieve an earlier diagnosis. As early detection improves clinical outcome, a number of screening tests including clinical and self-breast examinations, mammography, ultrasound, magnetic resonance imaging and genetic screening have been employed (Khatcheressian et al. 2006; Dey 2014). In a clinical setting, the diagnostic procedure is initiated with a clinical breast examination, mammography, and fine needle aspiration (FNA) or core/excisional biopsy, where a portion of the tumour or the entire lump is removed for tissue or cellular analysis.

Clinical breast examinations (CBE) every 3 years as well as frequent self-examinations have been recommended for women younger than 40 years (at average risk for breast cancer). For women over the age of 40 years however, annual CBEs have been suggested. In terms of

imaging, the most widely recommended screening approach in the United States for this group has been annual mammography. Mammography is a special type of low-dose x-ray imaging used to create detailed images of the breast. This screening method is currently the best available population-based method to detect breast cancer at an early stage, when treatment is most effective (Saarenmaa et al. 1999). The two types of mammography examinations include screening mammography, which is performed in asymptomatic women and diagnostic mammography, which is performed in symptomatic women when a breast lump, nipple discharge or abnormality is found during self-examination or screening mammography. Although mammography remains the most cost-effective approach for breast cancer screening, its sensitivity ranges between 69% and 90% (Rankin 2000) which slightly improves when combined with CBE.

Ultrasonography is generally used to assist the clinical examination of a suspicious lesion detected on mammography or physical examination. This technique is however limited by a number of factors including failure to detect micro-calcifications and poor specificity (33%) (Kuhl et al. 2000). Magnetic resonance imaging (MRI) has been explored as a modality for detecting breast cancer in women at high risk and in younger women. A combination of MRI techniques has been found to be highly sensitive to malignant changes in the breast when combined with mammography and CBE (~99%). The 10-fold higher cost and substantially poorer specificity (26%) compared to mammography resulted in limited use MRI screening as a general screening tool (Obdeijn et al. 2010). MRI screening has however been recommended in a subset of high-risk patients including those with BRCA mutations and a positive family history of cancer (Kriege et al. 2004; Warner et al. 2004).

Prognosis is an estimated prediction of outcome and the probability of progression-free survival (PFS) or disease-free survival (DFS) (American Cancer Society 2015). Survival is often calculated as an average number of months (or years) that 50% of patients survive, or alternatively, the percentage of patients that are alive after 1, 5, 15 and 20 years. Patients with good prognoses are typically offered less invasive treatments such as hormone therapy or lumpectomy and radiation, while patients with poor prognoses are offered more aggressive treatments, such as mastectomy and chemotherapy (Horwich et al. 1994). Prognostic factors include age of the patient, staging and disease recurrence. Stage is the most important prognostic factor as it takes into consideration tumour size, local involvement, lymph node status and the presence of metastatic disease (Table 2.2, summarized from Garcia et al. 2007 and http://www.breastcancer.org/symptoms/diagnosis/staging).

Table 2.2. Stages of Breast Cancer

Stage	Definition
Stage 0	Cancer cells remain inside the breast duct, without invasion into normal adjacent breast tissue.
Stage 1	Size of 2 cm or less and is confined to the breast (lymph nodes are clear).
Stage 2A	No tumour can be found in the breast, but cancer cells are found in the axillary lymph nodes (the lymph nodes under the arm).
	The tumour measures 2 cm or smaller and has spread to the axillary lymph nodes.
	The tumour is larger than 2 cm but no larger than 5 cm and has not spread to the axillary lymph nodes.
Stage 2B	The tumour is larger than 2 cm but no larger than 5 cm and has spread to the axillary lymph nodes.
	The tumour is larger than 5cm but has not spread to the axillary lymph nodes.
Stage 3A	No tumour is found in the breast; cancer is found in axillary lymph nodes that are sticking together or to other structures, or cancer may be found in lymph nodes near the breast bone.
	The tumour is any size; cancer has spread to the axillary lymph nodes, which are sticking together or to other structures, or cancer may be found in lymph nodes near the breast bone.
Stage 3B	The tumour may be any size and has spread to the chest wall and/or skin of the breast and,
	May have spread to axillary lymph nodes that are clumped together or sticking to other structures or cancer may have spread to lymph nodes near the breastbone.
	Inflammatory breast cancer is considered at least stage 3B.
Stage 3C	There may either be no sign of cancer in the breast or a tumour may be any size and may have disseminated to the chest wall and/or the skin of the breast as well as the lymph nodes (above or below the collarbone an axillary nodes near breast bone
Stage 4	The cancer has metastasized to other parts of the body.

A higher stage at diagnosis is typically associated with a worse prognosis. The greater the resemblance to normal cells, the slower the growth of the tumour cell and the better the prognosis. Grade on the other hand is a measure of the degree of differentiation of tumour cells. Cells that are not well differentiated will appear immature, divide more rapidly, and exhibit tendencies toward metastases. Well differentiated cells are assigned a grade of 1, moderate

grade 2, and poor or undifferentiated cells are allocated a higher grade of 3 or 4 (depending upon the scale employed). Early-onset disease tend to have a poorer prognosis due to a probable familial component linked to a more aggressive disease subtype such as estrogen receptor (ER), progesterone receptor (PR)-negative and human epidermal growth factor receptor-2 (HER2)–negative breast cancer (Fredholm et al. 2009; Partridge et al. 2010; Freedman and Partridge 2013).

2.7. Molecular profiling and Intrinsic disease subtypes

Molecular profiling of breast tumours, based upon variations in gene expression, has been used to characterize breast cancers beyond the conventional evaluation of grade, histology, and immunohistochemical (IHC) analysis of hormone receptors and HER2 overexpression. Breast cancer is among the earliest and most intensely studied diseases using DNA microarray platforms for gene expression profiling. This molecular profiling has enriched our understanding of disease heterogeneity and yielded valuable prognostic and predictive information. Four different gene lists including a varying number of genes have been used for molecular subtype identification and classification of breast cancer (Perou et al. 2000; Mackay et al. 2011). The resulting molecular classification has highlighted the existence of four major subtypes that overlap with different clinico-pathological classification systems (Su et al. 2011; Al-Ejeh et al. 2011).

The routine assessment of hormone receptor status is considered the standard of care in all newly diagnosed breast cancer patients. IHC and in-situ hybridization (ISH) techniques are used to determine ER, PR and HER2 receptor expression in early-stage breast cancer to prioritize the need for targeted treatment interventions. Variability in results reported has highlighted IHC/ISH testing limitations resulting from the lack of assay standardization, inter-observer variability in interpretation, as well as discrepancies between guidelines used to define cut-off limits for ER and HER2-positivitiy (Thomson et al. 2001; Regitnig et al. 2002). The development of RNA-based gene expression profiling techniques using microarray analysis and reverse transcription polymerase chain reaction (PCR) resulted from efforts to increase the prognostic and predictive value of the aforementioned conventional, routine tests. Grant et al. (2015) determined HER2 status across three assay platforms in an attempt to address the clinical dilemma presented by equivocal and contradictory results frequently obtained in patients with early-stage breast cancer. These authors demonstrated 100% concordance of HER2 status following IHC/FISH reflex testing of the identified discordant cases using microarrays (TargetPrint), thereby improving quality assurance and consequently the level of confidence on which treatment decisions were based.

Gene expression profiling technologies have allowed for the description of distinct breast cancer subtypes based on intrinsic molecular profiles (Perou et al. 2000; Prat and Perou 2011). It has been demonstrated that ER status is the main discriminator of molecular signatures which supports the long-standing hypothesis that ER-positive and ER-negative breast cancers are different diseases (Gruvberger et al. 2001). The different subtypes among ER-positive and ER-negative tumours have various implications for prognosis and treatment (Perou et al. 2000; Sotiriou et al. 2003). These observations suggest that gene expression profiling has the potential to change current breast cancer management, as evidenced with the use of the 70-gene MammaPrint profile (Agendia, Inc., Irvine, CA) in the South African population (Grant et al. 2013). MammaPrint uses a microarray platform for classification of early-stage breast cancer into low- or high-risk groups for chemotherapy selection. This test also enables sub-classification of tumour tissue into luminal A, luminal B, HER2-positive and basal-like subtypes (BluePrint), based on the ~25 000 genes in the human genome evaluated during development of the MammaPrint test (Tian et al. 2010).

At least four distinct breast cancer subtypes identified have recently been extended to about 10 subtypes with the use of next generation sequencing (NGS) (Curtis et al. 2012). Surrogate definitions of the intrinsic subtypes include two main subtypes of ER-negative tumours: basal-like and HER2-enriched, and three subtypes of ER-positive tumours: luminal A and luminal B (HER2-positive) and luminal B (HER2-negative).

2.7.1. Luminal A and B subtypes

Luminal cancers derive their name from similarity to genes expressed by normal breast luminal epithelial cells. They typically express luminal cytokeratins 8 and 18 and are the most common subtypes that make up the majority of ER-positive breast cancer. They are characterized by the expression of ER, PR, and other genes associated with ER activation.

Luminal A tumours make up ~40% of all breast cancers and are associated with the most favourable outcomes compared to other subtypes. They usually have high expression of ER-related genes, low expression of the HER2 cluster of genes, and low expression of proliferation-related genes. Luminal B tumours on the other hand are less common (~20%) and are subdivided into those expressing HER2 and those who lack HER2 expression. They exhibit relatively lower (than luminal A) expression of ER-related genes, variable expression of the HER2 cluster, and higher expression of the proliferation-related genes (Yersal and Barutca 2014; Ades et al. 2014).

Cell proliferation markers used to distinguish between luminal A and luminal B tumours (HER2-negative) present substantial challenges in the clinical setting. In cases where reliable marker indexes are not available, alternative measures of proliferation such as a histological grade are used in distinguishing between luminal A and B subtypes. More recently, microarray-based molecular profiling using the 80-gene BluePrint has been shown to accurately discriminate between the luminal A and B subtypes (Glück et al. 2013; Whitworth et al. 2014), the latter being characterized by increased tumour proliferation, higher recurrence rates, poorer overall prognosis (Correa Geyer and Reis-Filho 2009; Wirapati et al. 2008) and are compatible with a high-risk 70-gene prognostic profile (Leo et al. 2012; Grant et al. 2013). BluePrint molecular subtyping determines the mRNA levels of 80 genes that discriminate between the three breast cancer subtypes based on functional molecular pathways: Luminal, HER2, and Basal. MammaPrint substratifies Luminal into Luminal A (MammaPrint Low Risk) and Luminal B (MammaPrint High Risk). The BluePrint reclassified 22% of breast tumours, reassigning more responsive patients to the HER2 and Basal categories while reassigning less responsive patients to the Luminal category. Based on these findings, Whitworth et al. (2014) suggested that compared with IHC/FISH, BluePrint more accurately identified patients likely to respond (or not respond) to neoadjuvant chemotherapy.

2.7.2. HER2-enriched subtype

The HER2-enriched subtype (previously the ER-/HER2+ subtype) makes up about 10-15% of breast cancers and is characterized by high expression of HER2 and proliferation gene clusters, and low expression of the luminal cluster. For this reason, these tumours are typically negative for ER and PR, and positive for HER2. This subtype comprises only ~50% of clinically HER2-positive breast cancer (Eroles et al. 2012). The other approximate 50% display high expression of both the HER2 and luminal gene clusters and fall within the luminal B subtype. HER2-positive tumours are associated with increased cell proliferation, angiogenesis, tumour invasiveness, and a high nuclear grade (Chikarmane et al. 2015) and prior to HER2-targeted therapy, this subtype carried a very poor prognosis (Slamon et al. 2001; Cadoo et al. 2013).

The most appropriate management plan for a patient diagnosed with cancer may be informed by an assessment of the components of the patient's germline genome and their genome. In addition to the use of RNA/protein expression to distinguish between different breast cancer subtypes, single nucleotide polymorphisms (SNPs) have been investigated for their role in defining these subtypes. O'Brien et al. (2013) demonstrated that low-penetrance genetic risk factors for breast cancer vary by subtype and further clarified the role of several key susceptibility genes. The discovery of the involvement of SNPs in the determination of disease

subtype led to the development of high-density SNP arrays. Hansen et al. (2015) sought to determine whether examination of HER2 copy number using high-density SNP arrays can provide additional diagnostic power to assess HER2 gene status. They concluded that determination of HER2 copy number variations by SNP array-based genomic segmentation analysis provides additional diagnostic sensitivity and accuracy, thereby providing an effective supplement to IHC/FISH HER2 analysis so that more women may be elected for targeted treatment with HER2 inhibitors. SNP pharmacogenetic analysis to determine the need for trastuzumab treatment of HER2-positive patients may however on its own be insufficient to replace gene expression analysis and IHC. The knowledge and new insights gained through SNP analysis may furthermore identify targets for intervention prior to the development of cancer, serving as a tool for prevention which is the direction and ultimate goal of genetic testing. In line with the findings of Hansen et al. (2015), Kotze et al. (2015) suggested a comprehensive approach that combines tumour pathology, genetic and relevant biochemisty for the classification of patients into different treatment groups.

Variations in the HER2 gene have been identified that modify the protein sequence of the HER2-neu protein, but how these polymorphisms affect prognosis and response to HER2 targeted therapy is unknown. Su et al. (2015) genotyped 3 SNPs in the HER2 gene in 303 breast cancer patients to study their association with HER2 protein expression in breast cancer. Two SNPs (rs1058808 and rs2517956) were significantly associated with HER2 protein expression ($p=0.007$ and $p=0.008$, respectively), but not with other parameters such as clinical stage, tumour size, histological grade, lymph node metastasis, ER, PR, Ki67 and P53 status ($p>0.05$). SNPs in the recently discovered metastasis-associated in colon cancer-1 (MACC1) as well as in the Insulin-like growth factor 1 (IGF-1) and Vascular endothelial growth factor A (VEGFA) genes have been associated with the poor clinical outcome of HER2-positive breast cancer (Maae et al. 2012; Muendlein et al. 2013; Muendlein et al. 2014). In a case-control study, Stanton et al. (2015) studied 11 SNPs in relation to trastuzumab cardiotoxicity in 140 breast cancer patients. The HER2 Pro 1170 Ala SNP was significantly associated with trastuzumab cardiac toxicity ($p=0.04$) after adjusting for relevant confounders. This SNP may therefore identify a subset of breast cancer patients who are at increased risk of trastuzumab therapy cardiotoxicity and may be useful in conjunction with other biomarkers for stratifying patients into different risk categories for optimal clinical management.

2.7.3. Basal-like subtype

The basal-like subtype, named after its similarity in expression to that of the basal epithelial cells, makes up about 15-20% of breast cancers. It is characterized by low expression of the

luminal and HER2 gene clusters and is typically ER, PR, and HER2-negative, which has led to the term "triple-negative". While most triple-negative breast cancers (TNBCs) are basal-like, and most basal-like tumours are triple-negative, there is about 30% discordance between these two classifications (Bertucci et al. 2008). These tumours have high expression of the proliferation cluster of genes with exceptionally high mitotic rates, are virtually always high grade and are characterized by widespread genomic instability (even early in the disease) (Livasy et al. 2006; Fulford et al. 2006; Turner et al. 2007). Furthermore, they display high expression of the epidermal growth factor receptor, as well as a unique cluster of genes called the basal cluster, which include basal epithelial cytokeratins 5, 14, and 17. The basal-like subtype is strongly associated with BRCA1 mutation-positive cancers of which over 80% fall into this subtype. In addition, population-based studies suggest that this subtype is overrepresented in early-onset cases and in women of African-American descent (Carey et al. 2006; Morris et al. 2007; Bauer et al. 2007). Basal-like breast cancers are associated with poor prognosis and the worst survival rates compared to the luminal A and luminal B subtypes (Foulkes et al. 2003). Commercially available NGS tests (e.g. OncoDEEP, Belgium) may be particularly useful in triple-negative breast cancer or metastatic cancer (e.g. breast, colon, lung, ovarian, etc.) to determine functional pathways for the application of personalized medicine. Personalized medicine approaches use targeted therapies in subgroups of patients considered to derive the most benefit based on their genetic background.

The findings of Rummel et al. (2013) supported that of Fostira et al. (2012) who recommended that patients with early-onset TNBC should be considered eligible for BRCA1 mutation testing even in the absence of a family history of cancer. Fostira et al. (2012) demonstrated that 23% of BRCA1 mutation carriers had no reported family history of breast or related cancers, which implies that 23% of BRCA-mutation positive patients would have been missed had they not been tested on account of the absence of a positive family history.

Knowledge of the extent to which BRCA founder mutations, modifier genes and lifestyle risk factors relevant to the South African population correlate with hormone receptor status represents an important research question to ultimately guide treatment decisions. An increased risk of TNBC was observed among females with a first-degree family history of breast cancer (Phipps et al. 2011). While women with basal-like/triple-negative disease experience high short-term mortality (O'Brien et al. 2010, Perez et al. 2011), the development and FDA approval of targeted anti-HER2 therapy has improved survival rates for women with HER2-positive disease.

A frequently occurring phenomenon seen in breast cancer is altered tumour receptor status throughout disease progression - possibly influenced by adjuvant therapies - which significantly influences patient survival. Women with ER-positive primary tumours that switched to ER-

negative tumours had a 48% increased risk of mortality (95% CI, 1.08 to 2.05) compared with women with stable ER-positive tumours (Lindström et al. 2012). Accurate documentation of ethnicity, family history, mutation status, HR status and other relevant clinical information therefore becomes increasingly important for present and future classification of patients into different treatment groups.

2.8. Breast cancer treatment strategies

Breast cancer is treated by surgery, radiation, hormone therapy, chemotherapy and immunotherapy. The presence of estrogen and progesterone receptors on tumour cells is important in terms of guiding treatment. Two-thirds of invasive breast cancers are ER and PR-positive (Nicholson et al. 2003; Gonzales-Anguloetal et al. 2007) and require the hormones estrogen and progesterone to proliferate. A greater incidence of ER-positive tumours has been reported among postmenopausal women. Naturally, breast cancer patients who test negative for ER and PR will not respond to hormone therapy, bestowing their chances of survival on remaining treatment options (mainly chemotherapy), as well as disease morphology (Garcia et al 2007).

2.8.1. Surgery

The radical mastectomy, designed by William Stewart Halsted in 1889, was coined the Halsted radical mastectomy. About 90% of women in the United States of America (USA) who were treated for breast cancer underwent Halsted radical mastectomies from 1894 to about 1979. Although this procedure reduced the recurrence rate from 51-82% to 6% (Yang et al. 2005), it was a disfiguring surgery which increased morbidity to a great extent. Its consequences ranged from universal lymphedema (condition in which excess lymph fluid collects in tissues, causing swelling) to severely restricted arm movement, resulting in eventual chronic pain (Zurrida et al. 2011). Within that period, women who were diagnosed with breast carcinoma immediately underwent radical mastectomies in the belief that it was unequivocally necessary in order to prevent metastasis.

Currently, the various types of breast cancer surgery include:

1. Lumpectomy (breast-conserving surgery) that involves removal of the tumour and some unaffected tissue around it. The lymph nodes may also be removed by making a second incision under the armpit.
2. Partial/segmental mastectomy or quadrantectomy involving the removal of the cancerous area as well as the surrounding area of healthy tissue, more so than with a lumpectomy. This is usually followed by administration of radiation.

3. Simple or total mastectomy where the entire breast is removed, yet sparing the lymph nodes. Simple mastectomy is most frequently performed for prevention of new cancer development or when current cancer is in its early stages as in the case of in situ, micro-invasive, and stage IA breast cancers. With a double mastectomy however, both breasts are removed as a pre-emptive strategy to prevent development of cancer in the other breast in high-risk cases.
4. Radical mastectomy involves removal of the entire breast and nipple, lymph nodes in the armpit, and chest muscles under the breast. This procedure is currently rarely performed as a modified radical mastectomy is as effective and less disfiguring (The 1979 National institute for health consensus statement concluded that survival for modified radical mastectomy was equivalent to radical mastectomy).
5. Modified radical mastectomy whereby the entire breast along with the nipple is removed. Some lymph nodes in the armpits are removed, however the chest muscles are not removed.

2.8.2. Radiation therapy

Radiation therapy involves treatment with high-energy rays or particles purposed to destroy cancer cells. It may be administered externally (external beam radiation) or internally (brachytherapy) subsequent to breast-conserving surgery to reduce the risk of recurrence in the breast or lymph nodes (Vinh-Hung and Verschraegen 2004), or after mastectomy in patients with tumours >5 cm or metastasis to the lymph nodes and/or other areas of the body (Fisher et al. 1985; Fisher et al. 1989). A functional polymorphism in the Apurinic endonuclease 1 (APEX1) gene (rs1130409) involved in the base excision repair pathway has been associated with higher sensitivity to ionizing radiation and increased risk of treatment-related pneumonitis (Yin et al. 2011). The same variant has been shown to modulate prognosis of advanced non-small cell lung cancer in patients following platinum-based chemotherapy (Zhao et al. 2013).

2.8.3. Hormone treatment

The female sex hormones estrogen and progesterone are naturally produced by the ovaries and promotes the growth and survival of normal as well as cancerous breast epithelial cells by binding to and subsequently activating the ER. Estrogen-dependent processes that are important in the development and progression of the majority of breast cancers may be disturbed by the following 3 ways: 1) interference with the binding of estrogen to the ER and/or to the promoter elements of the genes it regulates by selective estrogen receptor modulators (SERMs) such as raloxifene and tamoxifen; 2) reduction or inhibition of ER expression by

fulvestrant, a selective ER down-regulator which acts by limiting ER availability for estrogen binding, and 3) reduction of the amount of estrogen produced via ovarian removal in premenopausal women and use of aromatase inhibitors/ inactivators (AIs) in postmenopausal women.

2.8.3.1. Aromatase inhibitor treatment

Aromatase inhibitors and inactivators act by hindering the body's ability to produce estrogen from androgens through the suppression of aromatase enzyme activity. While ovarian aromatase is responsible for the majority of circulating estrogen in premenopausal women, aromatase in fat and muscle may be responsible for much of the circulating estrogen in postmenopausal women. In tissues that are highly sensitive to estrogen including the breast, vagina, uterus, heart and blood vessels, bone and the brain, aromatase provides local estrogen in an autocrine manner. Consequently, breast aromatase activity is frequently increased in breast cancer (Simpson et al. 2001).

Three generations of AIs have been developed to date with each successive generation having a higher specificity for the aromatase enzyme, greater suppression of aromatase activity and fewer adverse events, and (Dowsett 1990; MacNeill et al. 1992; van der Wall et al. 1993; Dowsett et al. 1995; Geisler et al. 1996; Geisler et al. 1998). The use of first- and second-generation AIs was limited by side effects such as dizziness, fatigue, nausea and vomiting, rash, ataxia as well as by a lack of enzyme selectivity. Third-generation AIs are superior compared to their predecessors as they are associated with greater suppression of aromatase activity and fewer adverse events. The two classes of third-generation AIs include non-steroidal AIs (anastrozole and letrozole) which reversibly bind to the aromatase enzyme and steroidal AIs such as exemestane which irreversibly binds to aromatase. They are therefore currently the preferred treatment for ER-positive and/or PR-positive breast cancer in postmenopausal women in the early and metastatic setting (Winer et al. 2004). Adverse events include hot flushes, fatigue, vaginal dryness, loss of libido, joint stiffness, arthralgias and loss of bone mineral density with subsequent increased risk of fracture (Winer 2004). In premenopausal women, AIs have a limited ability to reduce circulating estrogen and are generally not prescribed in this subgroup of breast cancer patients without concomitant drugs that suppress the rise in gonadotrophins and subsequent increase in hormone levels (Winer 2005).

2.8.3.2. Tamoxifen treatment

Tamoxifen is a selective estrogen receptor modulator (SERM) that has been used for more than three decades as an endocrine treatment for HR-positive breast cancer, with indications in the

preventative, adjuvant and metastatic settings (Bernard et al. 2006). As an adjuvant therapy in early breast cancer it improves overall survival, and its widespread use is thought to have made a significant contribution to the reduction in breast cancer mortality (Ring et al. 2004). Used widely in the treatment of advanced breast cancer, it remains the standard adjuvant therapy for pre-menopausal women and a valid treatment option, alongside AIs for post-menopausal women (Newman et al 2008; Ferraldeschi et al. 2010).

Tamoxifen is a non-steroidal agent with potent anti-estrogenic properties which competes with estrogen for binding sites in breast and other tissues. It competitively binds to estrogen receptors on tumours and other tissue targets, thereby producing a nuclear complex that reduces DNA synthesis and inhibits the effects of estrogen (Ring et al. 2004; Ferraldeschi et al. 2010). It causes (pre)cancerous cells to remain in the G0 and G1 (quiescent) phases of the cell cycle, preventing cells from dividing rather than causing cell death. Its mechanism of action is therefore cytostatic as opposed to cytocidal (Lippman et al. 1999).

Tamoxifen is metabolised in the liver by the cytochrome P450 isoforms CYP2D6 and CYP3A4 into active metabolites such as 4-hydroxytamoxifen and N-desmethyl-4-hydroxytamoxifen (endoxifen), which have 30-100 times greater affinity to the estrogen receptor than tamoxifen itself (Figure 2.2). Since endoxifen is 30-100 fold more efficient as an estrogen antagonist than tamoxifen (Jordan et al. 1977), its plasma concentration is crucial in determining the efficacy the drug. Concentrations of active metabolites are affected by factors such as environmental drug-induced (eg. antidepressants) and genetic factors including CYP2D6.

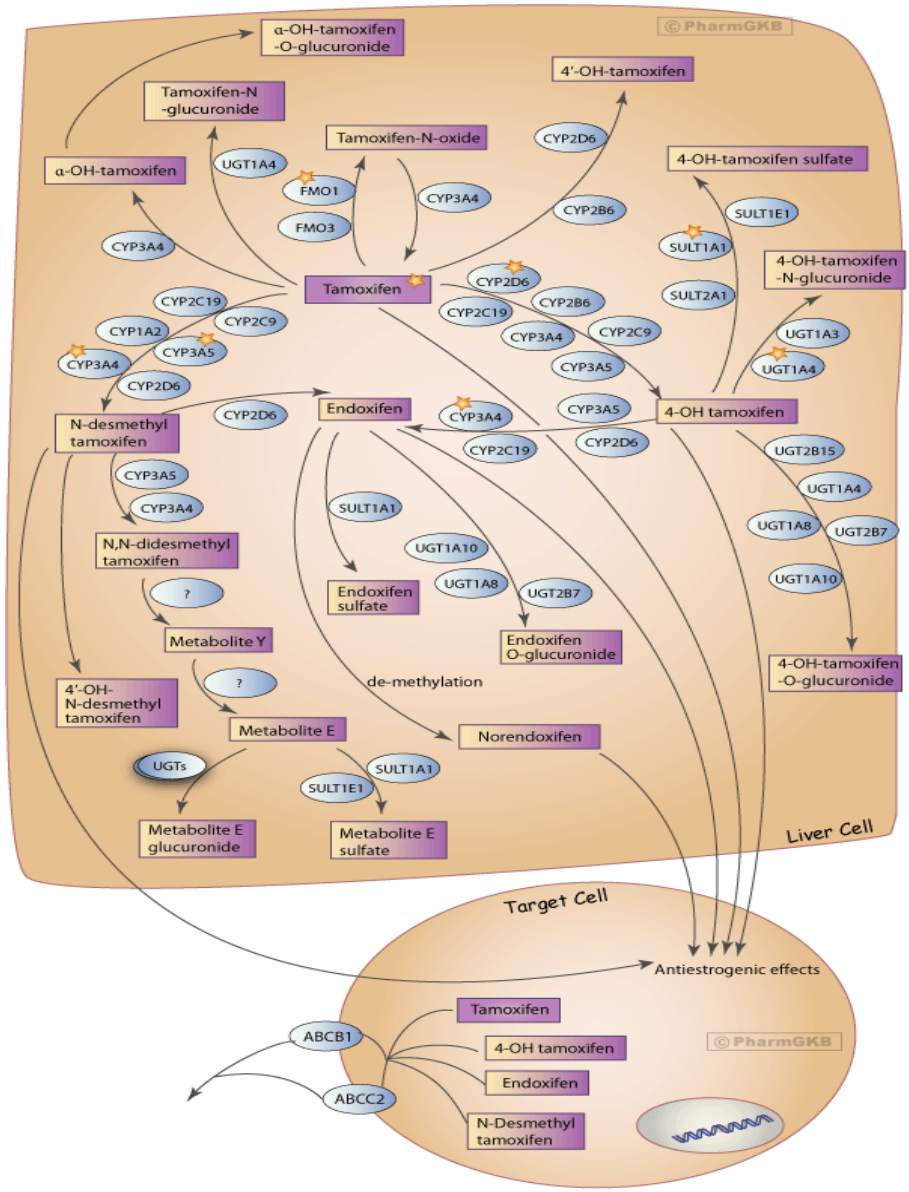


Figure 2.2. Tamoxifen metabolism in the liver (reproduced with permission from PharmKB, https://www.pharmgkb.org/pathway/PA145011119).

Although tamoxifen has several beneficial effects, approximately 30% of all patients treated with this drug do not experience the full medical benefit due to the co-administration of other drugs or genetic mutations in the gene encoding the enzyme that bio-transforms it into its active metabolite (Ferraldeschi and Newman 2010). This leads to significant inter-individual variability in drug metabolism and differences in drug response, making the beneficial and adverse effects of tamoxifen virtually unpredictable for individual patients before the availability of genetic testing (Bernard et al. 2006).

Tamoxifen binds to the ER which in turn interacts with DNA. The tamoxifen/ER complex recruits other proteins known as co-repressors to inhibit estrogen-mediated gene activation. Some of these proteins include the nuclear receptor co-repressor (NCoR) and the silencing mediator for retinoid and thyroid receptor (SMRT) (Shang et al. 2000; Ring et al. 2004). The effectiveness of tamoxifen can be influenced by various elements including growth factors. High levels of growth factor proteins including ErbB2/HER2 have been shown to occur in tamoxifen-resistant cancers (Jansen et al. 2005). Tamoxifen seems to require protein PAX2 for its full anti-cancer effect since the tamoxifen/ER complex is able in the presence of high PAX2 expression to suppress the expression of the proliferative HER2 protein (Hurtado et al. 2008). In contrast, when AIB-1 expression is higher than PAX2, the tamoxifen/ER complex up-regulates the expression of HER2, resulting in stimulation of breast cancer proliferation.

Tamoxifen through its antagonistic effect on the ER has been shown to reduce the risk of bone loss in postmenopausal women in addition to exerting favourable effects on osteoporosis prevention in postmenopausal women (Nakamura et al. 2007). In contrast, tamoxifen appears to be associated with bone loss in pre-menopausal women, particularly those who continue to menstruate after adjuvant chemotherapy (Vehmanenet et al. 2006). The dual role of tamoxifen is evidenced by its antagonistic effects in breast tissue as well as its partial agonistic effect on the endometrium. The rationale for using tamoxifen for a limited period of 5 years is therefore based on the substantial increased risk for endometrial cancer (Grilli et al. 2006).

A significant number of tamoxifen-treated breast cancer patients experienced a reduction of libido (Cella et al. 2006). Although tamoxifen treatment in post-menopausal women was associated with beneficial effects on serum lipid profiles (de Medina et al. 2004), long-term data from clinical trials have failed to demonstrate a cardio-protective effect (Esteva and Hortobagyi 2006). In women with variation in the apolipoprotein E (ApoE) gene, tamoxifen can however cause a rapid increase in triglyceride concentration in the blood (Liberopoulos et al. 2002). Tamoxifen use has also been associated with increased risk of thromboembolic events (TEs) in women at high risk for the disease due to variation in genes involved in haemostasis including factor V leiden (FVL) (Garber et al. 2010), prothrombin (FII) and MTHFR (Elhassan and Abdalla 2015). The risk of thromboembolism with tamoxifen use is also increased during and immediately after major surgery or periods of immobility (Decensi et al. 2005). Tamoxifen is also a contributory cause of fatty liver, otherwise known as hepatosteatosis (Osman et al. 2007). Tamoxifen-treated breast cancer patients showed evidence of reduced cognition (Paganini-Hill et al. 2000) and semantic memory scores (Eberling et al. 2004), which may be exacerbated in carriers of the MTHFR 677 C>T and ApoE E4 mutations (Lengacher et al. 2015). A study by

Bender et al. (2001) demonstrated that memory impairment in patients treated with tamoxifen was less severe compared with those treated with the AI Anastrozole.

Tamoxifen has been described as a chemo-preventative agent, as it decreased the incidence of invasive and non-invasive breast cancer. Studies involving more than 50 000 tamoxifen-treated women have produced prevention or risk reduction results in the range of 43% to 49% in three distinct invasive breast cancer risk settings of 1) healthy high-risk women, 2) patients with early-stage breast cancer (reduction of contralateral breast cancer risk), and 3) patients with DCIS (Fisher et al. 1998; Fisher et al. 2005). These preventive effects of tamoxifen highlight the remarkable consistency of this treatment agent in settings of high breast cancer risk (Lippman et al. 1999). According to the Early Breast Cancer Trialists' Collaborative Group (2005), allocation to about 5 years of adjuvant tamoxifen treatment (for ER-positive disease only) reduced the annual breast cancer death rate by 31%, regardless of the use of chemotherapy, age (<50, 50-69, ≥70 years), PR status or other tumour characteristics. They contended that 5 years of tamoxifen treatment is significantly more effective than just 1-2 years of the treatment ($p<0.001$ for recurrence, $p=0.01$ for breast cancer mortality). It has also been suggested that tamoxifen may reduce breast cancer risk through mechanisms other than receptor-mediated estrogen inhibition (Ring and Dowsett 2004).

The lack of hormone receptor positivity in BRCA1 tumours suggests that treatment with tamoxifen, or other hormonal therapy, may be less effective in these patients. BRCA1 tumours are mostly ER and PR-negative while BRCA2 tumours are more commonly positive for both receptors (Karp et al. 1997; Lomen et al. 1998). TNBCs are common among patients with BRCA1 mutations (Lakhani et al. 2005). The finding that TNBC is most common among premenopausal women, especially those of African descent (Lund et al. 2009; Morris et al. 2007), emphasizes the importance of the development of novel chemotherapeutic drugs such as poly ADP ribose polymerase (PARP) inhibitors that target BRCA-deficient breast cancer.

2.8.4. Chemotherapy

Chemotherapy predominantly used for stage 2-4 breast cancers is particularly beneficial in ER-negative disease (Hortobagyi 1998; Miller and Sledge 1999). The chemotherapy agents are often administered in combinations for periods of 3 to 6 months. Chemotherapy agents work by destroying fast-growing and/or fast-replicating cancer cells, mainly by causing DNA damage and interference of DNA replication. These agents may however also damage normal cells, which may cause severe side effects including cardiac toxicity (Bird and Swain 2008), contralateral and secondary non-breast cancer (Schaapveld et al. 2008; Beadle et al. 2009), sexual and reproductive effects (Partridge et al. 2008), neurotoxicity and cognitive impairment (Brezden et

al. 2000; Bender et al. 2001; Ahles et al. 2002; Falleti et al. 2005; Vardy 2009) and anxiety and depression (Silva et al. 2010; Hack et al. 2012). The most commonly used neoadjuvant chemotherapy regimens contain an anthracycline (adriamycin or epirubicin) (AC) in combination or sequentially administered with taxanes (paclitaxel or docetaxel) (CAT). Anthracyclines can be combined with cyclophosphamide and fluoropyrimidine. Another common combination treatment includes cyclophosphamide, methotrexate (MTX) and 5-fluorouracil (5FU) (CMF).

The folate pathway is an important target for both cancer prevention and treatment. The anti-cancer activities of CMF used as the cornerstone chemotherapeutic agents for treatment of metastatic disease are exerted through inhibition of this pathway (Shrubsole et al. 2005). It is therefore noteworthy that a number of studies have associated the 677 C>T mutation in the MTHFR gene with increased clinical response to 5FU, but an increased toxicity to MTX (Yood et al. 2012; Chiusolo et al. 2002; Toffoli et al. 2003; Cohen et al. 2003). Concomitant folate supplementation is generally recommended during treatment with MTX to reduce toxicity of MTX. While high folate intake was associated with a decreased risk of postmenopausal breast cancer in some studies (Ericson et al 2007; Lajous et al. 2006), others could not confirm this association (Feigelson et al. 2003; Rohan et al. 2000). Conflicting results may be attributed to the dual role of folate in cancer development: folic acid supplementation may provide protection early in carcinogenesis, yet may promote carcinogenesis if administered later and at very high intakes (Ulrich 2006; Kim 2006). This may pose a clinical dilemma in patients with two copies of the MTHFR T-allele, who have an increased requirement for folate above the daily recommended dose (RDA) (Moriyama et al. 2002). Since folic acid (synthetic folate) is more bioavailable than folate obtained from natural sources, single-nutrient high-dose supplementation is not recommended for cancer patients (Ulrich 2007). Synthetic folic acid requires several reduction reactions to produce the active molecule used in one carbon metabolism and therefore sufficient intake as part of a healthy diet should be encouraged or supplementation with an active form of folate recommended. Early detection of MTHFR 677 TT homozygotes as a high-risk genetic subgroup with increased requirements of folate and other vitamin co-factors is essential to optimize the enzymatic functional balance between DNA synthesis, repair and methylation processes (Friso et al. 2002; Peerbooms et al. 2011; Cohen et al. 2012).

2.8.4.1. Platinum-based chemotherapy

The platinum salts (cisplatin, carboplatin, and oxaliplatin) exert their anti-cancer activity by directly binding to DNA, causing the formation of DNA-platinum adducts and subsequently intra- and inter-strand DNA crosslinks. These DNA crosslinks ultimately impede replication and

transcription, leading to single and double-strand DNA breaks with resultant cytotoxicity (Noll et al. 2006).

Cisplatin (cis- diamminedichloroplatinum), also known as Peyrone's chloride, is one of the most widely used chemotherapeutic drugs. It was approved in 1978 as the first drug in its class, and early on was demonstrated to be active in the first-line treatment of advanced breast cancer (Sledge et al. 1998). Over time and with the availability of better-tolerated therapeutics, platinums fell out of favor in the treatment of advanced breast cancer. While cisplatin has a strong anti-cancer activity, it also exerts negative side effects including toxicity to the kidneys and nervous system (Kelland 2007). Carboplatin is a bidentate carboxylate ligand approved by the FDA in 1989 for the treatment of ovarian, head and neck, and lung tumours (Dasari and Tchounwou 2014). The DNA adducts formed by this molecule are the same ones introduced by cisplatin and thrombocytopenia is its main negative side effect.

The last platinum drug approved by FDA in 2002 is oxaliplatin. With a slightly different mechanism of action than the aforementioned platinums, oxaliplatin is effective in the treatment of carboplatin and cisplatin-resistant tumours (Raymond et al. 2002) and in combination with 5-fluorouracil (5-FU), is successfully used to treat colorectal cancer (Kelland 2007).

The more recently developed phenanthriplatin is one of the most promising drugs as this compound kills cancer cells more efficiently than cisplatin and oxaliplatin and appears to be immune to acquired resistance mechanisms (Park et al. 2012).

2.8.4.2. PARP inhibitors

Poly (ADP-ribose) polymerase (PARP) inhibitors represent a new and promising class of chemotherapeutic drugs aimed at disrupting PARP function in homologous recombination-defective cancers (O'Shaughnessy et al. 2009). PARP is a family of proteins with enzymatic properties, scaffolding properties, and recruiting ability for other necessary DNA repair proteins. PARP1 and PARP 2 are the best known of these proteins, with PARP1 being the most expressed member of the family and playing a critical role in base excision repair (BER) (Rouleau et al. 2010). PARP1 inhibition was found to be effective in the treatment of BRCA1 and BRCA2-positive cancers. Its effectiveness is based on the ability to induce single-strand breaks which leads to stalling of replication forks as well as the formation of double-strand breaks that cannot be repaired in the absence of functional BRCA1 and/or BRCA2 proteins. This results in high levels of genomic instability and ultimately cell death. Olaparib was the first PARP1 inhibitor to be approved by the US food and drug administration (FDA) for the treatment of ovarian cancers with BRCA mutations. Currently, other PARP1 inhibitors including niraparib,

iniparib, veliparib, rucaparib and BMN-673 are being assessed in clinical trials, alone or in combination with chemotherapy or radiotherapy, for several sporadic tumours (Tangutoori et al. 2015).

2.8.5. HER2 Immunotherapy

In addition to the assessment of HER2 oncogene amplification and/or overexpression as a prognostic marker, testing is used to determine eligibility for HER2-targeted therapies including trastuzumab, lapatinib and pertuzumab used in clinical practice. Trastuzumab has been shown to reduce recurrence risk by up to 50% when combined with chemotherapy (Piccart-Gebhart et al. 2005; Romond et al. 2005). Despite its great potential for risk reduction, indications for trastuzumab treatment in relation to tumour size and nodal involvement are controversial (National Comprehensive Cancer Network practice guidelines 2009, Park et al. 2010). A tumour size of <1 cm is considered small enough to obviate the need for chemotherapy (Grant et al. 2014). In addition, quantitative scoring of the HER2 gene copy number rather than HER2:-CEP17 ratio to define HER2 positivity is supported by Hanna et al. (2014). A HER2 gene copy number of <2.2 is considered ineffective for classification as HER2-positive.

Trastuzumab is a humanized recombinant monoclonal antibody directed specifically against HER2 receptor. It is currently used as part of the neoadjuvant treatment regimen in patients with HER2-positive breast cancer (Kaufmann et al. 2007). Trastuzumab is the first developed agent targeting the HER2 pathway and its binding to the extracellular domain of HER2 receptor leads to inhibition of tumour cell growth. Mechanisms of anti-tumour action include antibody-dependent cell-mediated cytotoxicity, inhibition of cleavage of the extracellular domain of the HER2 receptor, inhibition of ligand-independent HER2 receptor dimerization, inhibition of downstream signaling pathways and angiogenesis, induction of cell-cycle arrest and apoptosis, and interference with DNA repair (Molina et al. 2001; Baselga et al. 2001; Spetor et al. 2009).

Lapatinib is a small, orally active molecule which reversibly inhibits HER-1 and HER2 tyrosine kinase. This inhibition leads to blockage of the MAPK and PI3/Akt signaling pathways, resulting in growth arrest and/or apoptosis (Xia et al. 2002). Some data indicate that lapatinib can also block HER2-HER3-mediated cell growth (Amin et al. 2010; Ghosh et al. 2011). Being a small molecule that can penetrate the blood-brain barrier, it has also been studied in the treatment and prevention of brain metastases (Saleem et al. 2015; Tomasello et al. 2010).

Pertuzumab is the humanized monoclonal antibody that binds to dimerization domain II of HER2 receptor, which is necessary for HER2 activation and cell signaling. Clinically, the most

important action of pertuzumab is inhibition of HER2-HER3 dimerization. Pertuzumab affects important signaling pathways that mediate cell proliferation and synergistically with trastuzumab inhibits breast tumour cell survival (Graus-Porta et al. 1997; Citri et al. 2003; Baselga and Swain 2009).

2.8.6. mTOR inhibitors

Aberrations in the phosphatidylinositol-3-kinase/protein kinase B/mammalian target of rapamycin (PI3K-AKT-mTOR) pathway are frequently observed in breast cancer. This pathway plays an important role in the regulation of cell proliferation, metabolism, motility, angiogenesis, and survival. Pathway hyperactivation has been linked to cancer pathogenesis, progression, and treatment resistance (Shaw et al. 2006). Cancer cells develop resistance to various therapies through activation of the PI3K-AKT-mTOR pathway, hence the addition of mTOR inhibitors to endocrine therapy, chemotherapy, and anti-HER2 therapy to enhance efficacy and/or delay treatment resistance have been considered. Dysregulation resulting in hyperactivation of the mTOR pathway has been associated with a poor outcome in breast cancer (Jerusalem et al. 2014).

2.9. Genetic contribution to breast cancer development and recurrence

The impact of inherited genetic variations associated with adverse clinical outcomes in breast cancer can be divided into distinct categories of risk based on their penetrance and prevalence in the general population (Coate et al. 2010) (Tables 2.3 to 2.5). Highly penetrant mutations of low prevalence mainly in the BRCA1 and BRCA2 tumour suppressor genes are responsible for early-onset familial breast cancer in 5-10% of all affected patients across ethnic groups. A study in non-BRCA families identified mutations in several well-known, high and moderate susceptibility genes including ataxia telangiectasia mutated (ATM) (~5%), the double strand break repair protein (RAD50) (~3%), checkpoint kinase 2 (CHEK2) (~ 2%), partner and localizer of BRCA2 (PALB2) (~1%), meiotic recombination 11, homologue A (MRE11A) (~1%) and tumor protein p53 (TP53) (~1%) (Aloraifi et al. 2015). Novel pathogenic variants in 30 other genes could collectively explain the missing heritability in up to 35% of BRCA-negative patients.

Table 2.3. High penetrance breast cancer predisposing genes

Gene name	Protein function	Reference
High penetrance genes		
CDH1	Cadherin-1 (CDH1) is a tumour suppressor gene and is important in cell–cell adhesion in epithelial tissue. Germline mutations in CDH1 predispose to diffuse gastric cancer and early onset breast cancer, specifically associated with lobular carcinoma.	Dossus and Benusiglio 2015; Nelen et al. 1996; Berx et al. 1995; Lalloo and Evans 2012; http://www.uniprot.org/
MLH1	MLH1 is a component of the post-replicative DNA mismatch repair system. It is also implicated in DNA damage signaling which induces cell-cycle arrest and may lead to apoptosis in case of major DNA damages.	Bianchi et al. 2011; Hamann 2000; http://www.uniprot.org/
MSH2	MSH2 is a component of the post-replicative DNA mismatch repair system. It forms two different heterodimers which binds to DNA mismatches thereby initiating DNA repair.	Westenend et al. 2005; Hamann 2000; http://www.uniprot.org/
PTEN	PTEN facilitates regulating of cell growth. A mutated PTEN gene causes Cowden syndrome, a rare disorder in which people have higher risks of both benign and cancerous breast tumours, as well as growths in the digestive tract, thyroid, uterus, and ovaries.	Hamann 2000; Nelen et a. 1996; Berx et al. 1995; Lalloo and Evans 2012; http://www.uniprot.org/
STK11	STK11 plays a role in the suppression of cellular proliferation and apoptosis. Gene mutations confer a risk of various cancers including female breast cancer.	Hamann 2000; Manolio et al. 2009; Hemminki et al. 1998; http://www.uniprot.org/
TP53	TP53 regulates the expression of target genes, thereby inducing cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism. Mutations in this gene are associated with a variety of human cancers including hereditary cancers.	Hamann 2000; Dossus and Benusiglio 2015; Manolio et al. 2009; Hemminki et al. 1998; http://www.uniprot.org/

Table 2.4. Moderate penetrance breast cancer predisposing genes

Gene name	Protein function	Reference
Moderate penetrance genes		
ATM	ATM helps repair damaged DNA and has been linked to an increased rate of breast cancer in some families as inheriting one defective copy of the gene prevents adequate repair of damaged DNA.	Hollestelle et al. 2010; Broeks et al. 2000; Swift et al. 1987; http://www.uniprot.org/
BARD1	BARD1 coordinates a range of cellular pathways including DNA damage repair, ubiquitination and transcriptional regulation to maintain genomic stability. It plays a central role in cell cycle control in response to DNA damage.	Seal et al. 2006; Ripperger et al. 2008; http://www.uniprot.org/
BRIP1	BRIP1 is required for the maintenance of chromosomal stability. It is also involved in the repair of DNA double-strand breaks by homologous recombination in a manner that depends on its association with BRCA1.	Seal et al. 2006; Hollestelle et al. 2010; http://www.uniprot.org/
CHEK2	CHEK2 is a cell-cycle checkpoint regulator and putative tumour suppressor. When activated, the encoded protein stabilizes the p53 ptotein, leading to cell cycle arrest. It also interacts with and phosphorylates BRCA1, allowing it to restore survival after DNA damage. Gene mutations confer risk for sarcomas, breast and brain cancers.	Hollestelle et al. 2010; Kainu et al. 2000; Meijers-Heijboer et al. 2002; http://www.uniprot.org/
MRE11	MRE11 is involved in DNA double-strand break repair and possess single-strand endonuclease activity and double-strand-specific 3'-5' exonuclease activity. It is also involved in meiotic double-strand break processing.	Shibata et al. 2014; Bartkova et al. 2008; http://www.uniprot.org/
MUC1	MUC1 is an oncoprotein that is aberrantly overexpressed in breast cancer cells and associates with HER2 at the surface of breast	Siroy et al. 2013; Raina et al. 2014; http://www.uniprot.org/

	cancer cells.	
NBS1	NBS1 is a member of the MRE11/RAD50 double-strand break repair complex. It is involved in DNA double-strand break repair and DNA damage-induced checkpoint activation.	Bogdanova et al. 2008; Nseir et al. 2005; Hollestelle et al. 2010; http://www.uniprot.org/
PALB2	PALB2 interacts with BRCA2 and allows it to connect with BRCA1, forming the BRCA complex. Mutations in this gene disrupt DNA repair pathways	Tischkowitz et al. 2010; Hollestelle et al. 2010; Rahman et al. 2007; http://www.uniprot.org/
RAD50	RAD50 is essential for double-stranded DNA break repair interacting with BRCA1, NBN and MRE11. An abnormal RAD50 gene has been linked to a higher risk of breast cancer in some families as cells cannot sufficiently repair damaged DNA.	Shibata et al. 2014; Damiola et al. 2014; http://www.uniprot.org/
RAD51	RAD51 participates in homologous recombination and double-strand break repair processes. It binds to both single and double-stranded DNA and forms part of a PALB2-scaffolded homologous recombination complex containing BRCA2 and RAD51C. It also plays a role in regulating mitochondrial DNA copy number under conditions of oxidative stress in the presence of RAD51C and XRCC3.	Park et al. 2008; Sigurdsson et al. 2002; Sage et al. 2010; Pellegrini et al. 2002; http://www.uniprot.org/

Table 2.5. Low penetrance breast cancer predisposing genes

Gene name	Protein function	Reference
Low penetrance genes		
APEX1	APEX1 is a multifunctional protein that plays a central role in the cellular response to oxidative stress. The two major activities of APEX1 is DNA repair and redox regulation of transcriptional factors.	Hsu et al. 2014; Hinz et al. 2014; Chohan et al. 2015; http://www.uniprot.org/
CASP8	Most upstream protease of the activation cascade of caspases are responsible for the TNFRSF6/FAS mediated and TNFRSF1A induced cell death.	http://www.uniprot.org/
FGFR2	FGFR2 is a tyrosine-protein kinase that acts as cell-surface receptor and plays an essential role in the regulation of embryonic development, cell proliferation, differentiation, migration and apoptosis.	Fletcher et al. 2013; Robbez-Masson et al. 2013; http://www.uniprot.org/
H19	The H19 gene product is a long non-coding RNA which functions as a tumour suppressor. Mutations in this gene have been associated with Beckwith-Wiedemann Syndrome and Wilms tumourigenesis.	Medrzycki et al. 2014; Yan et al. 2015; http://www.uniprot.org/
HRAS1	HRAS1 belongs to the Ras oncogene family, and products encoded function in signal transduction pathways. Defects in this gene are implicated in a variety of cancers, including bladder cancer, follicular thyroid cancer, and oral squamous cell carcinoma.	http://www.uniprot.org/
LSP1	LSP1 encodes an intracellular F-actin binding protein which is expressed in lymphocytes, neutrophils, macrophages, and endothelium. It may regulate neutrophil motility, adhesion to fibrinogen matrix proteins, and transendothelial migration.	http://www.uniprot.org/

Table 2.5. Low penetrance breast cancer predisposing genes

Gene name	Protein function	Reference
Low penetrance genes		
MAP3K1	MAP3K1 is a component of a protein kinase signal transduction cascade. It activates the ERK and JNK kinase pathways by phosphorylation of MAP2K1 and MAP2K4 as well as CHUK and IKBKB, the central protein kinases of the NF-kappa-B pathway.	Hu et al. 2014; Glubb et al. 2015; http://www.uniprot.org/
RPAP1	RPAP1 forms an interface between the RNA polymerase II enzyme and chaperone/ scaffolding protein. It is also required for interaction of the RNA polymerase II complex with acetylated histone H3.	Jeronimo et al. 2007; http://www.uniprot.org/
TNRC19	The TNRC19 encoded protein is a member of the membrane-associated guanylate kinase homologue (MAGUK) family that participates in the assembly of multi-protein complexes on the inner surface of the plasma membrane at regions of cell-cell contact.	http://www.uniprot.org/

Highly penetrant mutations of low prevalence are responsible for early-onset familial breast cancer in 5-10% of all affected patients with 3-4 % resulting from mutations in the BRCA1 and BRCA2 tumour suppressor genes, across ethnic groups (García et al. 2004). This led to the development of diagnostic BRCA mutation screening assays and computer-based algorithms for risk assessment routinely applied in clinical practice worldwide, including South Africa (Schoeman et al. 2013). The high lifetime risk for developing familial breast cancer estimated at 65% to 85% for BRCA1 and 45% to 85% for BRCA2 mutation carriers (Chen et al. 2007), led to the development and implementation of personalized screening programs, prophylactic surgery and/or chemo-preventive therapies in mutation-positive patients and their at-risk family members (Schrag et al. 2000; Paradiso and Formenti 2011). It is estimated that known breast cancer susceptibility loci account for less than 50% of the familial risk of the disease, with the remaining variance likely to be due to genetic variants conferring low-moderate risks (Easton et al 2007). The tumour suppressor genes BRCA1 and BRCA2 are involved in maintenance of genome stability, specifically the homologous recombination pathway for double-strand DNA

repair. Guidelines for diagnostic and susceptibility screening for high penetrance mutations such as those in the BRCA1 and BRCA2 genes are well established and carried out routinely in clinical practice worldwide. Diagnostic BRCA testing has been available in South Africa since 2004, and has been facilitated by the identification of common founder mutations in the BRCA1 and BRCA2 genes in Afrikaner and other local populations (Reeves et al. 2004, van der Merwe et al. 2012b). Knowledge of hormone receptor status - as determined by transcriptional profiling to select South African patients for chemotherapy (Grant et al. 2013) and anti-HER2 treatment (Grant et al. 2015) - is also relevant in relation to familial breast cancer since women with the triple-negative and basal-like subtypes are at increased risk for carrying BRCA1 mutations (Lakhani et al. 2005).

Over the preceding decade increasing evidence has been provided in numerous studies showing that multifactorial diseases may result from a combined effect of multiple low-penetrance alleles and environmental factors (Pagon et al. 2002). These include functional gene variants underlying shared metabolic pathways that are potential nutritional and drug targets across diagnostic boundaries (Li et al. 2009; Ehrnhoefer et al. 2011). Polymorphic variation in the MTHFR and CYP2D6 genes are important modifiers of breast cancer development and recurrence, respectively, in BRCA mutation carriers (Jakubowska et al. 2007; Pepe et al. 2007; Newman et al. 2008). In an effort to optimize strategies for effective recurrence risk prediction or prevention of cumulative risk in patients with familial breast cancer, we proposed that CYP2D6 genotyping is indicated in breast cancer patients who 1) are receiving tamoxifen and 2) are at high risk for tumour recurrence due to a family history of breast cancer/BRCA-mutation positive or 3) are required to take potential competing antidepressants.

2.9.1. Cytochrome P450 2D6 (CYP2D6) pharmacogenetics

CYP2D6 is one of the best studied cytochrome P450 enzymes, with numerous variant alleles designated *1 to *61. CYP2D6 plays an important role in the oxidative metabolism of various drugs such as neuroleptics, tricyclic antidepressants (TCA's) and selective serotonin reuptake inhibitors (SSRI's), adrenergic-blocking drugs (metoprolol), anti-arrhythmic drugs (sparteine and propafenone), opioids (codeine) (Zanger et al. 2004; Lohmann et al. 2001) and cholesterol-lowering statins (Li et al. 2014). Nearly 99% of poor metabolisers have any two of the following alleles: *3, *4, *5, *6, *7 *8 or *11 (Bosch et al. 2006; Bradford et al. 2002). CYP2D6 catalyzes the conversion of tamoxifen to more potent metabolites, and poor CYP2D6 enzymatic activity has been associated with tamoxifen treatment outcome (Bosch et al. 2006).

The CYP2D6 gene is located on chromosome 22q13.1 and more than 60 different alleles have been linked to this enzyme (Zanger et al. 2008; Zanger et al. 2004). The CYP2D6*4 allele

(rs3892097) is the most common variant and results in a stop codon as a direct result of a change in the splice site acceptor site (slicing defect, G to A base change at position 1846 in intron 3). Imparting increased risk of adverse reactions from drugs requiring CYP2D6 metabolism for activation or elimination, 5-10% of individuals in the Caucasian population have been shown to inherit the allele 4 null-allele (i.e. poor metabolisers), compared to 1-3% in other ethnic groups. Individuals who inherit two copies of the CYP2D6*4 which accounts for more than 75% of poor metabolisers in Caucasians (Bradford et al. 2002) may be more susceptible to hypersensitivity, toxicity or a lack of response to certain drugs (Griese et al. 1998; Zanger et al. 2004).

Therefore, alterations in CYP2D6 activity can negatively affect treatment outcomes in patients receiving antidepressant pharmacotherapies that are CYP2D6 substrates (D'Empaire et al. 2011). Poor metabolisers have two copies of a defective CYP2D6 gene, and as a result they metabolise drugs at a much slower rate. This may result in a greater potential for adverse events and lower efficacy of drugs requiring CYP2D6 activation (Bernard et al. 2006). Cronin-Fenton and Lash (2011) suggested that the effects of CYP2D6 inhibition on the response to tamoxifen treatment are likely small or non-existent. In this review, the authors concluded that extended CYP2D6 genotyping as well as the assessment of other genes related to tamoxifen metabolism and clearance should be considered an important avenue for future research. Abraham et al. (2010) also concluded that, while the CYP2D6*6 genotype could play a role in breast cancer survival in tamoxifen-treated patients, that the relationship between the inactivating CYP2D6*4 allele and therapeutic outcomes in relation to endocrine treatment remains incompletely described and warrants further investigation.

A substantial body of evidence implicates genetic variation in the CYP2D6 and CYP2C19 enzymes, which catalyze the hydroxylation and demethylation of many drugs and neurochemical compounds, in the response to antidepressants and other psychiatric medications. It is estimated that establishing CYP2D6 metabolizer status through extensive genotyping for drug-inactivating alleles will benefit treatment with 30-40% of drugs metabolised, equating to nearly 10% of those currently in use. Reduced enzymatic activity associated with polymorphisms in the CYP2D6 gene is of particular clinical relevance in the context of combination therapy with agents such as certain SSRI's that may act as inhibitors of this enzyme. A greater understanding of a patient's phenotype and metabolizer status as established by genotype has significant clinical implications in many non-communicable diseases, particularly in cancer patients using concomitant medications that are substrates and inhibitors for CYP2D6.

Findings from a study by Sestak et al. (2012) failed to support the role of CYP2D6 genotype as a determinant of adverse clinical outcomes or the development of side effects related to endocrine treatment in patients with breast cancer. Brauch and Schroth (2012) reviewed the relationship between CYP2D6 genotype and response to tamoxifen treatment. The authors concluded that, given the limitations and potential for study bias inherent to prospective-retrospective studies, a prospective trial would be necessary to resolve issues related to this clinical question. Ongoing trials in the metastatic setting are not considered sufficient to resolve this issue, although hormone-insensitive patients are eligible for the Eastern Cooperative Oncology Group E3108 (ECOG E3108) clinical trial based on tamoxifen citrate use in patients with metastatic or recurrent breast cancer (Brauch et al. 2013). This notion is also relevant to results reported following the Breast International Group (BIG) 1-98 CYP2D6 study, in which bias introduced due to the use of tumour as opposed to host DNA resulted in massive deviation of the CYP2D6 genotype distribution from Hardy-Weinberg equilibrium (Nakamura et al. 2012). This led to a call to retract published findings related to this trial.

The discovery that relatively frequent SNPs in genes coding drug-metabolising enzymes or those dependent on vitamin co-factors for optimal activity can explain differences in treatment response, drug toxicity and/or recurrence risk, support the development of genetic tests used for selection of the most appropriate treatment strategy for each patient (Del Re et al. 2012). In addition to CYP2D6, variation in the genes encoding MTHFR for example, are important modifiers of breast cancer development and recurrence in BRCA mutation carriers (Jakubowska et al. 2007; Pepe et al. 2007; Newman et al. 2008). Knowledge of such risk-modifying factors in addition to BRCA mutation status may aid in effective recurrence risk prediction or prevention strategies in patients with familial breast cancer.

2.9.2. MTHFR nutrigenetics: folate metabolism and the methylation pathway

Folate is a nutrient involved in the methionine metabolic pathway, which is essential for DNA methylation and is also required for DNA synthesis (Stover 2004). Cumulative data suggests that abnormalities in folate metabolism facilitate cross-talk between genetic and epigenetic processes by affecting gene expression (DNA methylation) and genome integrity (DNA synthesis and repair) (Kim 1999; Choi and Mason 2002). Folate polyglutamates are converted to 5,10-methylenetetrahydrofolate (methylene-THF), which is required as a methyl donor by dUMP (2′-deoxyuridine-5′-monophosphate) for the synthesis of dTMP (2′-deoxythymidine-5′-monophosphate) (Figure 2.3).

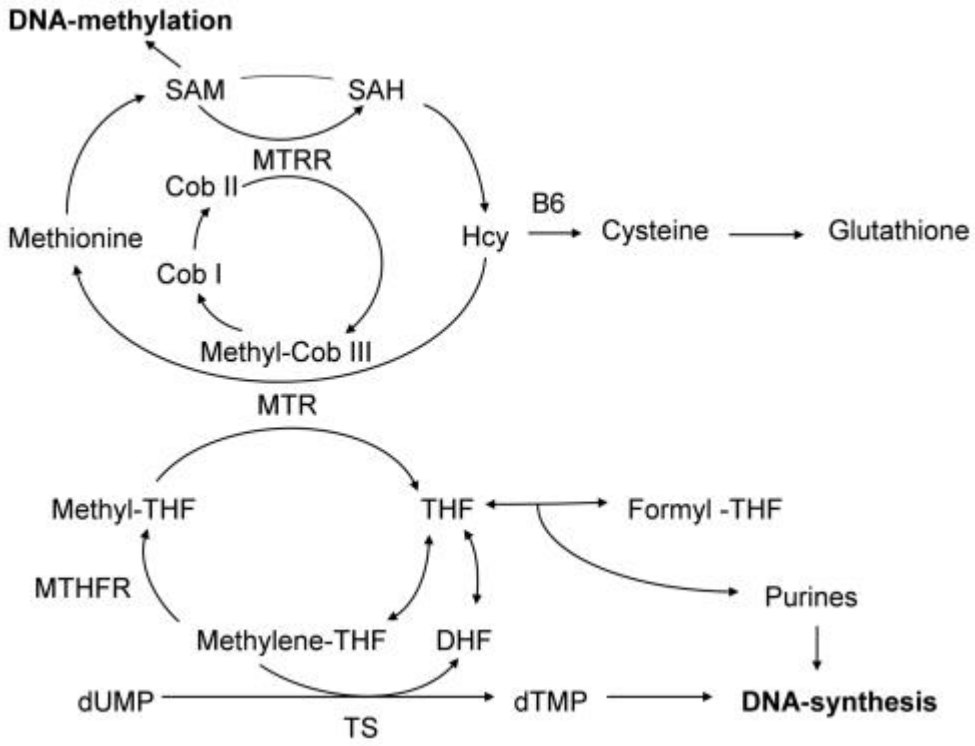


Figure 2.3. Overview showing the action of the MTHFR, MTR and MTRR enzymes in the folate and methionine metabolism pathways. TS, thymidylate synthase; DHF, dihydrofolate; Formyl-THF, 10-formyltetrahydrofolate; Cob I, cob(I)alamin; Cob II, cob(II)alamin; Methyl-Cob III, methyl cob(III)alamin; SAH, S-adenosylhomocysteine; Hcy, homocysteine; B6, vitamin B6. Figure reproduced with permission from Wettergren et al. (2010).

Methylene-THF is the precursor of metabolically active 5-methyltetrahydrofolate (methyl-THF) which is used in the remethylation of homocysteine (sulfur-containing amino acid) to methionine. Methionine is then broken down to produce the universal methyl donor S-adenosylmethionine (SAM). The conversion of methylene-THF to methyl-THF is dependent on the MTHFR enzyme. The gene encoding the MTHFR enzyme is located on chromosome 1, position p36.3 (Goyette et al. 1994) and contains 11 exons (Frosst et al. 1995). Variation in the MTHFR gene has been linked to reduced enzyme activity resulting in hyperhomocysteinemia when folate status is low.

The most widely studied variations in the MTHFR gene include two functional polymorphisms at nucleotide positions 677 (C>T, rs1801133) and 1298 (A>C, rs1801131). The point mutation 1298 A>C results in the substitution of alanine to glutamine at amino acid position 429 (Glu429Ala or E429A) of the regulatory domain of the enzyme (Rozen 1997; van der Put et al. 1998). MTHFR 677 C>T is caused by a cytosine to thymine change at nucleotide position 677 and results in a substitution of valine to alanine at amino acid position 222 (A222V) (Frosst et al. 1995). This gene variant is the most common genetic cause of hyperhomocysteinaemia and has

been linked to a reduction in activity and increase in thermoliability of the MTHFR enzyme (Jacques et al. 1996; Schwahn and Rozen 2001; Yamada et al. 2001). The MTHFR CT-heterozygous and TT-homozygous genotypes result in 65% and 30% reduction in enzyme activity, respectively (Tajouri et al. 2006). The frequency of the polymorphism varies significantly between ethnic groups (Schneider et al. 1998). Homozygosity for the MTHFR 677 TT genotype is higher than 20% in certain US populations such Hispanics, Columbians and Amerindians in Brazil, whereas in the Black population the frequency is less than 2%. The frequency for the Caucasian populations in countries such as Europe, North America and Australia is 8-20% (Gilbody et al. 2007) while the MTHFR 677TT genotype occurs in approximately 10% of individuals in the North American population. This is similar to the frequency reported in the South African Caucasian population, while this genotype was not detected in Black Africans studied by Scholtz et al. (2002). These authors reported an intermediate frequency in the South African Coloured population of mixed ancestry.

The 1298 A>C mutation is caused by an A to C substitution at nucleotide position 1298 in exon 7 (glutamine to alanine amino acid change) of the MTHFR gene. The frequency of the homozygous MTHFR 1298 CC genotype in Caucasian populations is 7-12% (North America and Europe). A reduced prevalence is observed in Hispanics, with a frequency of 4-5% compared to frequencies of 1-4% in the Chinese and Asian populations. Evidence exists for an association between this polymorphism and depression (Gilbody et al. 2007). We found that in depression patients but not in controls, the minor T-allele of MTHFR 677 C>T was associated with high BMI when folate intake was low, which in turn correlated significantly with increased homocysteine levels (Delport et al. 2014).

It is known that MTHFR 677 C>T causes reduced enzyme activity that may impair folate metabolism, resulting in genomic DNA hypomethylation (Ulrich et al. 2007) as well as gene-specific promoter hypermethylation (Oyama et al. 2004), as is frequently observed in cancer. In addition, reduced MTHFR activity may lead to increased plasma levels of homocysteine when folate status is low. Increased levels of homocysteine may in turn promote inflammatory processes by generating reactive oxygen species (ROS) (Mato et al. 2005) thereby playing a crucial role in cellular oxidative stress. Homocysteine may furthermore affect hypo- and hypermethylation of DNA (Peyrin-Biroulet et al. 2007). Evaluation of the MTHFR 677 C>T polymorphism as a predictive marker for colorectal cancer patients treated with 5-FU-based chemotherapy points to a link between the MTHFR 677 T allele and tumour response (Cohen et al. 2003; Etienne et al. 2004; Derwinger et al. 2009; Chua et al. 2009). Studies have also associated this polymorphism with 5-FU-related cytotoxicity (Derwinger et al. 2009; Chua et al. 2009).

The extent to which epigenetic lesions are induced by our genetic background, environment, or by interaction of the two resulting in combined effects remains uncertain. Germline variants in the MTHFR and MTR genes found to be associated with aberrant patterns in tumour pathology in patients with colorectal, breast and lung cancer provides an important example of the interplay between genetic and epigenetic factors in the cancer cell (Paz et al. 2002). Carriers of the MTHFR 677 T-allele display low levels of 5-methylcytosine in their tumour genomes and do not achieve severe degrees of global hypomethylation. Analysis of tumours found to be homozygous GG for the methionine synthase (MTR/MS) 2756 A>G variant furthermore showed a low number of hypermethylated CpG islands in tumour-suppressor genes. The 2756 A>G polymorphism leads to an aspartic acid to glycine change in the protein-binding region of the MTR gene (van der Put et al. 1997), which results in a less effective enzyme that causes a reduction in vitamin B12 levels (Leclerc et al. 1996; Harmon et al. 1999; Klerk et al. 2003).

Occasionally, MTR loses its activity through oxidation of cob(I)alamin to cob(II)alamin. The methionine synthase reductase (MTRR) enzyme then reactivates MTR in a reaction that generates methyl(III)cobalamin (Zijno et al. 2003). The MTRR 66 A>G polymorphism leads to a substitution of the amino acid isoleucine with that of methionine at codon 22 (Wilson et al. 1999) and results in an enzyme exhibiting a 4-fold reduction in activity compared with the normal protein *in vivo* (Olteanu et al. 2002). Thus, the MTRR 66 G-allele may decrease the availability of SAM by reducing the levels of active MTR, and as a result, induces DNA hypomethylation. The MTR 2756 GG genotype has been associated with an increased risk of colorectal cancer (de Vogel et al. 2009), whereas a reduced risk of colorectal adenoma recurrence has been observed in patients possessing one copy of the MTRR 66 A>G variant (de Vogel et al. 2009). The effects of the MTR 2756 A>G and MTRR 66 A>G variants are influenced by intake of B-vitamins (Koushik et al. 2006; Hubner et al. 2006). Kang et al. (2005) recently showed that in uterine and cervical cancer patients with heterozygous genotypes with both of these gene variants, the average number of methylated genes was lower than in patients with the homozygous AA genotypes.

The importance of the activity of the aforementioned enzymes in the folate-methylation pathway makes them attractive candidates for analysis of functional genetic variants in carcinogenesis. Several studies confirmed that functional polymorphisms in the MTHFR gene modify the risk of breast cancer and may potentially alter the risk of ovarian cancer in women with an inherited predisposition (Chen et al. 2005; Jakubowska et al. 2007). The MTHFR 677TT genotype was associated with a 2-3 fold increased risk of breast and ovarian cancer among Polish BRCA1 mutation carriers when compared to individuals with the wild-type and heterozygous genotypes. Presence of the MTHFR 677 T-allele was found to confer an increased risk of breast cancer in

BRCA1 mutation carriers (p=0.007), while the presence of the 1298 C-allele confers an increased risk of breast cancer in sporadic cases (p=0.015) (Pepe et al. 2007). The MTHFR 677TT genotype has furthermore been linked with congenital abnormalities, spontaneous abortions/miscarriages, renal tube defects, acute leukemia, motor dysfunction, seizures and thrombosis (De Mattia and Toffoli 2009). Aberrant genomic DNA methylation as a consequence of reduced MTHFR enzyme activity may be caused by variation in the MTHFR gene and/or low folate status *(*Friso et al. 2002). Klerk et al. (2002) performed a meta-analysis in which individuals homozygous for the MTHFR 677 C>T polymorphism were found to have a 16% higher risk of coronary heart disease in comparison with those homozygous for the wild-type allele. According to Wald et al. (2002), lowering homocysteine concentrations by 3 mmol/l by increasing dietary folate intake, would reduce the risk of ischemic heart disease by 16%, deep vein thrombosis by 25%, and stroke by 24%.

MTHFR, MTR and MTRR implicated in the development and progression of cancer, CVD, depression and treatment-related side effects (Peerbooms et al. 2011) act as treatment targets across the disease spectrum due to their involvement in one-carbon metabolic pathway essential for DNA synthesis and the epigenetic process of DNA methylation.

2.10. Cancer Epigenetics

It is well recognized that epigenetic changes contribute to carcinogenesis and provide a mechanistic link between lifestyle (environmental and dietary) exposures and disease. Epigenetic literally means “in addition to changes in genetic sequence.” (Weinhold 2006). It refers to alterations in gene expression/activity without changing the DNA sequence. A consensus definition of the concept of epigenetics was formulated at a Cold Spring Harbor meeting in 2008, as "stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence". Some studies showed that these modifications can be transmitted to daughter cells whereas other experiments demonstrated that these changes may be reversed (Li et al. 2005; Iglesias-Linares et al. 2010; Wang et al. 2010). Many types of epigenetic processes have been identified including methylation, acetylation, phosphorylation, ubiquitylation, and sumolyation. With regard to carcinogenesis, these mechanisms exist within two overarching processes, namely 1) chromatin remodeling by histone methylation or histone acetylation, and 2) DNA methylation via repetitive elements, DNA methyltransferases (DNMTs), CpG islands and gene expression.

One of the most extensively studied epigenetic mechanisms is the formation of CpG sites by the pairing of cytosine with guanine by means of a phosphate group which has been a major focus of epigenetic research due to its capacity to directly silence gene expression, particularly with

respect to tumour-suppressor genes in carcinogenesis. This mechanism involves methylation of the fifth carbon of a cytosine nucleotide to create 5-methylcytosine (5mC), the methyl group of which lies in the major groove of the DNA helix that interferes with transcription factor binding to prevent gene expression. Moreover, there is a class of methylated DNA-binding proteins which binds to methylated cytosines and represses gene transcription by blocking transcription factors. DNA methylation is highly dysregulated in cancer. Aberrant patterns of methylation arise, leading to hypomethylation of distal regulatory regions and repetitive elements along with hypermethylation of CpG islands (Bird 2002; Ehrlich 2009). It has been known for some time that tumours from different sites display distinct CpG methylation profiles (Esteller et al. 2001) and exhibit distinct pathways of carcinogenesis within tumour sites (Sartor et al. 2011; Shen et al. 2007). Figure 2.4 illustrates the roles of DNA damage, epigenetic deficiencies in DNA repair and mutations in the pathogenesis and progression to cancer, while Figure 2.5 depicts common DNA damaging agents including examples of lesions they cause in DNA, together with the genes involved in the pathways used to repair these lesions. It also indicates the genes that are epigenetically regulated, and which of them display reduced expression in various cancers.

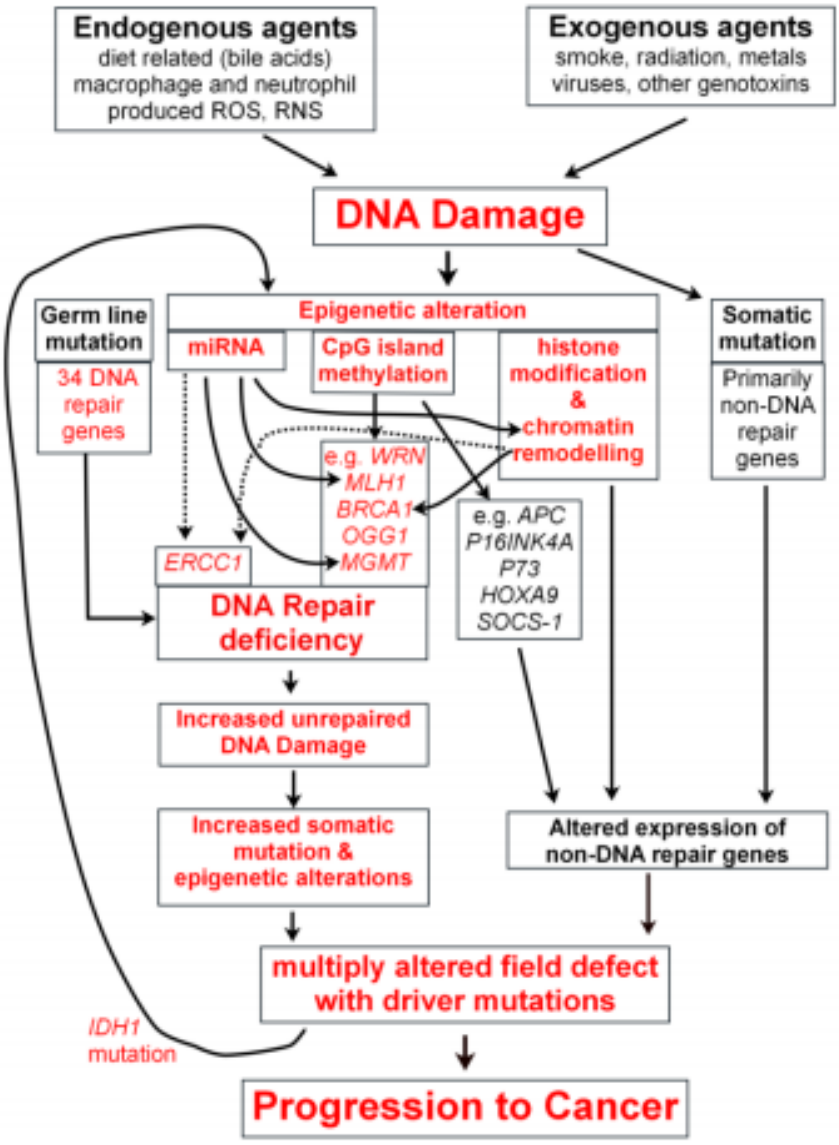


Figure 2.4. The roles of DNA damage, epigenetic deficiencies in DNA repair and mutation in progression to cancer [reproduced with permission from Bernstein et al. (2013a)].

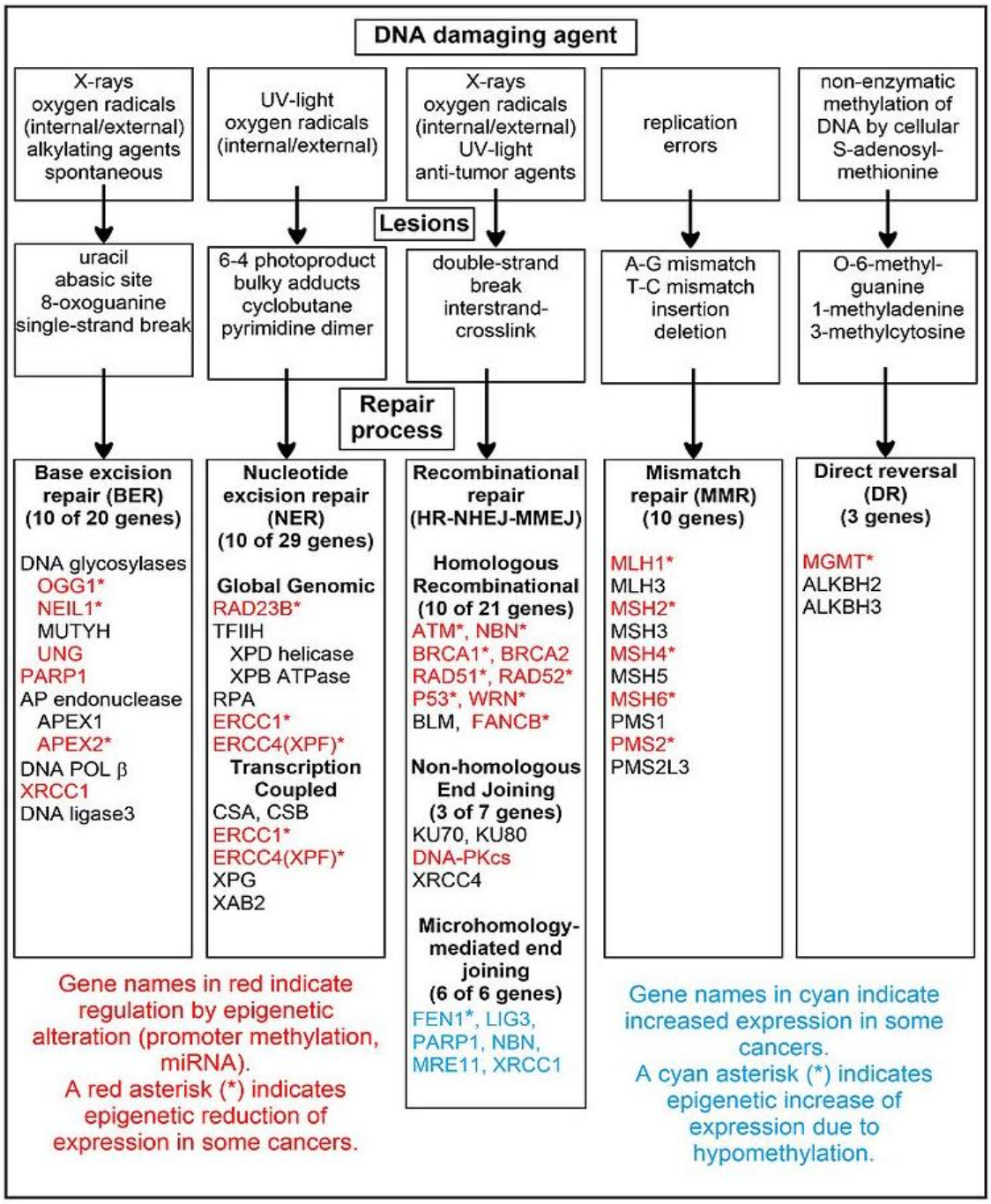


Figure 2.5. A chart of common DNA damaging agents including examples of lesions they cause in DNA, and genes involved in the pathways used to repair these lesions. Also indicated are the genes that are epigenetically regulated, and which of them display reduced expression in various cancers [reproduced with permission from Bernstein et al. (2013b)].

Variable penetrance of familial breast cancer mutations highlighted the limited predictive value of single-gene testing as in the case of BRCA1 and 2 that is modified by gene-gene and gene-environment interaction. Despite the success of linkage studies in identifying BRCA mutations in addition to other rare, causative mutations in single-gene disorders, a gap in knowledge was identified regarding limited power in detecting causative variants in complex diseases. In this context, the identification of moderate-high risk alleles other than those in the well-known BRCA genes may explain some of the “missing heritability”. Moreover, GWAS have identified several susceptibility variants which have been implicated as potential risk modifiers both in Mendelian and multifactorial diseases such as cancer. A number of common (mostly sporadic) breast cancer susceptibility loci have been associated with a slightly increased or decreased risk of breast cancer. To date, 74 SNPs have met the criteria for genome-wide “discovery” (Hindorff et al. 2013). These can act synergistically with environmental/lifestyle factors or follow the polygenic model to account for a small fraction of familial breast cancer cases. Most of these low-susceptibility loci have been highlighted through GWAS and initially included a number of loci. Although the precise contribution of low penetrance, common susceptibility loci in hereditary breast cancer is uncertain, the identification of such alleles may explain a subset of breast cancer cases. The identification of low-risk polymorphisms requires a paradigm shift in our understanding of the pieces of the susceptibility jigsaw, continuously being collected (Khoury et al. 2007). How they all fit together and what the implications are for genetic counselling and clinical care are currently being clarified with regard to interaction between different elements of disease susceptibility, both genetic and environmental. The overall importance of these polymorphisms remains high, because their effect appears to be multiplicative, such that an individual possessing several polymorphisms in DNA repair genes may have a significantly increased risk of breast cancer (Antoniou et al. 2008; Pharoah et al. 2008, Ricks-Santi et al. 2013).

The use of NGS technologies has to a large extent superseded GWAS and has become increasingly popular in discovering and validating putative gene variants associated with the pathogenesis of breast cancer and associated comorbidities.

2.11. Obesity and associated diseases

Obesity is an important determinant of cardiovascular risk, especially in younger age groups (Krauss et al.1998). Increased weight is a determinant of higher levels of triglycerides, elevated low-density lipoprotein (LDL)- cholesterol (C) and low high-density lipoprotein (HDL)-C. Obesity therefore has a great influence on lipoprotein metabolism, regardless of ethnic background. Results from numerous, large prospective cohort studies have generally showed an increased

risk of CVD mortality with increasing BMI (Wilson et al. 2002; McGee 2004; Flegal et al. 2005). Several studies have also associated obesity and weight gain with breast cancer risk (Huang et al. 1997; Wolket al. 2001; Stephenson and Rose 2003) and there is evidence that weight loss, as well as reduced fat intake may lead to decreased risk for breast cancer (Harvie et al. 2005; Prentice et al. 2006). In addition to cancer, obesity has been associated with various diseases listed in Table 2.6.

Table 2.6. Obesity and its associated diseases.

Disease	Prevalence/ risk
Cancers	Up to 20% of all cancer deaths are attributed to obesity (30% of endometrial cancers) (Calle et al. 2003; Chang et al. 2013)
Metabolic syndrome (MetS)	30% of middle-aged individuals in developed countries display symptoms of MetS (Després and Lemieux 2006)
Depression	Obesity associated with 55% increased risk of developing depression (Luppino et al. 2010)
Hypertension	5x risk amongst obese individuals; 66% of hypertensive cases linked to excess weight (Foresight report 2007)
Type 2 Diabetes	44% of type 2 diabetes attributed to obesity (World Health Organization 2013)
Coronary artery disease (CAD) and stroke	3.6x risk of CAD for each unit change in BMI and 6x risk of stroke in obese individuals (Prospective Studies Collaboration 2009)
Osteoarthritis (OA)	Increasing BMI frequently associated with OA in the elderly; risk of disability due to OA equal to CVD and greater than any other medical condition in the elderly (Grotle et al. 2008)
Reproductive function	6% of primary infertility in females attributable to obesity and impotency and infertility in men frequently associated with obesity (Hammoud et al. 2008; Arendas et al. 2008)
Respiratory effects	40% prevalence of obstructive sleep apnoea in obese individuals and up to 48% of patients with pulmonary hypertension were obese (Vgontzas et al. 1994; Taraseviciute and Voelkel 2006)
Gall bladder disease	3x risk of gall bladder disease in women with of BMI of >32 kg/m^2 and 7x risk with BMI >45 kg/m^2 (Bonfrate et al. 2014)
Liver disease	Up to 80% of obese patients have non-alcoholic fatty liver disease and up to 20% have non-alcoholic steatohepatitis (Sanyal et al. 2002; Bugianesi et al 2002)

Due to the known effects of obesity, alcohol and smoking on DNA methylation implicated in breast cancer and its associated co-morbidities, assessment of MTHFR as part of a comprehensive screening approach may facilitate the prevention of cumulative risk across diagnostic boundaries. Testing for MTHFR 677 C>T as an invariant marker to assess the risk for co-morbidities associated with dysfunctional methylation reflected by plasma homocysteine levels, provides added value compared to isolated biochemical analysis which is influenced by various environmental and lifestyle factors. For this reason, MTHFR is not only a feasible anti-cancer treatment target, but also a target for prevention of cumulative risk.

Although the clinical utility of single-gene testing in a high-risk population group is well-established, evaluation of multiple low-penetrance SNPs is complicated by uncertainty related to selection criteria and interpretation of the results for clinical application (Kotze et al. 2013). Pathology-supported genetic testing (PSGT) as a proposed solution was developed to overcome these limitations and may serve as a screening step for the selection of patients for whole exome sequencing (WES). Therefore, the CVD multi-gene assay previously described by Kotze and Thiart (2003) has been incorporated as the genetic component of a chronic disease screening program ideally suited to bridge the GWAS-related knowledge gap while moving from single-gene diagnostic testing to next generation sequencing (NGS) in genetically uncharacterized patients (Kotze et al. 2015). This approach may be particularly relevant in BRCA-negative breast cancer patients with certain comorbidities and/or experiencing therapeutic intolerance not adequately covered by the PSGT approach or accounted for by functional SNPs expressed in a high-risk environment.

2.12. Next Generation Sequencing

NGS applies to genome sequencing, genome resequencing, transcriptome profiling (RNA-Seq), DNA-protein interactions (ChIP-sequencing), and epigenome characterization (de Magalhães et al. 2010). The development of NGS technologies has escalated due to the demand for low-cost, high-throughput DNA sequencing which produces millions of sequences concurrently (Church 2006; Hall 2007; Schuster et al. 2008). For the translation of identified rare variants into clinical practice, larger sample size and greater ability to interpret the results to facilitate clinical diagnosis indicates exome sequencing may convey the most valuable method (Choi et al. 2009). As a means to check the usefulness of NGS as a tool for clinical testing, a targeted re-sequencing study of 21 hereditary breast cancer genes' exons were captured in subjects with a known mutation, which successfully detected 100% of mutations in all samples (Walsh et al. 2010).

Alterations to high-throughput sequencing methods have been developed depending on costs and efficiency needed by the study design: namely custom targeted, WES, or whole genome sequencing (WGS). Targeted sequencing allows selective capture of genomic regions of interest, either in-solution or by molecular inversion probes, which can then be amplified. Specifically for the Illumina platform, these sequences are subjected to the formation of clusters, which can be sequenced in parallel, in a single experiment. The sequence reads generated for each individual are mapped to a reference genome and variants can be identified and examined for pathogenicity. Using this technology, an enormous amount of data can be generated per region, thus providing great confidence for variant identification.

2.12.1. Whole Exome Sequencing

WES involves sequencing of all protein-coding regions (exons) in the genome and provides a cost-effective advancement for dissecting the genetic basis of complex diseases such as cancer and associated comorbidities. Compared with WGS, WES facilitates handling of data and generates higher-quality data, since it can be performed at greater sequencing depth. Various factors including the sample source, the specific region being sequenced and the particular algorithm used to compile and analyze the results determine the optimal sequencing depth. While exons represent only 1% of the genome, they account for approximately 85% of disease-causing variants, particularly for Mendelian traits (Bamshad 2011).

WES and WGS enable comparison of allele and genotype frequencies of previously published pathogenic rare variants between patient cohorts across ethnic groups (Norton et al. 2012). In addition, these sequencing methods can successfully aid clinical diagnosis (Choi et al. 2009; Gahl et al. 2012), reveal the genetic basis of rare familial diseases (Ng et al. 2010; Bamshad et al. 2011), and explicate novel disease biology (Lupski et al. 2010). This is of particular relevance to the elucidation of variants of “unknown clinical significance” (VUS) frequently identified in breast cancer patients subjected to NGS of the entire BRCA1 and 2 genes (Cheon et al. 2014), increasingly used to replace limited BRCA1/2 sequencing in familial cases.

A limitation of WES is that it predominantly identifies variants in the coding region of genes and is not able to detect disease-associated structural variation identifiable using other more comprehensive methods such as WGS (Ng et al. 2009). While the application of WES in the field of medical science is vast and highly advantageous, every step along the path to identifying potential disease-causing/contributory gene variants is complicated by technical, clinical and ethico-legal challenges.

Even when sequencing individual exomes, a large quantity of data and sequence information is generated and requires a significant amount of data analysis. Various challenges associated with the analysis of this data include differences in programs used to align and assemble sequence reads (Kahvejian et al. 2008; Dewey et al. 2011). Sequence technologies also have different error rates and generate varying read-lengths which may pose challenges when comparing results from different sequencing platforms. Furthermore, genetic heterogeneity due to ethnicity is a major limitation as it may increase the number of false positive and false negative findings in relation to the reference genome used, making the identification of candidate genes more difficult. The most striking technical challenges currently faced include the discrepancies between different sequencing platforms and low concordance among multiple variant-calling pipelines used (Boland et al. 2013). Low concordance among variant-calling pipelines limits the application of WES in clinical practice (O' Rawe et al. 2013).

Several tools have been developed to process exome sequence data. Generally, these tools are combined into a semi-automated workflow or pipeline, which include the following major steps.

2.12.2. Quality control of raw data

NGS platforms generate millions of sequence reads per run that each have a per-base quality score, which must be used as part of the analysis pipeline to identify problems that occurred during the generation of the data. This enables detection of problems that originated from the sequencer or library material used during the sequencing process, thereby preventing error from promulgating and producing false variant calls. It also provides an opportunity to salvage valuable data by clipping off the poor quality segments that generally occur at the 3'end of reads.

The most commonly used software for quality analysis is FASTQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc), which supports several input file formats produced by the various sequencing platforms. It can be run in two different modes, one as part of the sequence analysis pipeline producing an HTML QC report suitable for systemic processing of large numbers of datasets, or as an interactive application. Several analysis modules produce a range of quality metrics for sequence data. These modules range from basic statistics, such as overall nucleotide composition to various sequence quality scores. Systemic error is indicated when the quality score of the sequences is low overall and indicates poor sample quality/preparation or instrument failure. Contamination of the library or other issues during sequencing are also detected by an imbalance in the GC content at any read position, which should be very similar to the overall GC content of the underlying genome.

2.12.3. Mapping of reads to a standard reference genome

Sequence alignment is a critical step in exome sequencing analysis as it allows for mapping of the sequences to the reference genome, thereby detecting mismatches that represent SNPs, point mutations and insertion-deletion mutations (indels). Popular open source software tools used for this purpose are Tmap (https://github.com/iontorrent/TMAP), BWA (Li and Durbin 2009), Bowtie (Langmead et al. 2009) and BFAST (Homer et al. 2009). These tools display relatively equal performance and several are utilized in clinical sequencing applications. A shared shortcoming however, is their ability to manage genuine indels in short read data such as Illumina and SOLiD (Oliver 2012). A commercial aligner, NOVOALIGN (http://novocraft.com) has been shown to outperform the free software tools in this regard and recent releases have further improved its ability to accurately detect relatively large indels. A recent study has shown that it not only outperforms other popular aligners in the detection of indels, but also in the identification of SNPs close to indels (Pattnaik et al. 2012). Accuracy of the reference genome mapping step is paramount to identifying true variants.

2.12.4. Variant calling

In line with the third challenge identified in this study, accurate detection of genetic variants from mapped NGS reads is a crucial part of the WES analysis pipeline and determines the usefulness of a program's application in clinical practice. As even the best read mappers erroneously map a small fraction of reads, it is important that software for calling variants based on those alignments is able to detect and rectify mapping errors.

Several variant calling suites exist, including the Ion Torrent Variant Caller software plugin, the Ion Reporter's Ingenuity variant Caller (http://mendel.iontorrent.com/ion-docs/Ion-Reporter-Software-Features-Related-to-Variant-Calling.html), SAMtools (Li et al. 2009), the Genome Analysis Toolkit (GATK) (https://www.broadinstitute.org/gatk/index.php), CRISP (Bansal 2010), DiBayes (http://solidsoftwaretools.com/gf/project/dibayes/), BWT (Liand and Durbin 2009), VarScan (Koboldt et al. 2009) and SNPSeeker (Druley et al. 2009). GATK in particular is a multiple-sample, technology-aware single nucleotide variant caller. It is the most widely used variant calling toolkit (McKenna et al. 2010), which includes two modules called the GATK UnifiedGenotyper and HaplotypeCaller, the latter being superior in relation to indel detection. Software tools used to compensate for the presence of false-negative and false-positive variant calls include the RAREVATOR, a Perl script that executes the UnifiedGenotyper module of GATK, devised to filter and annotate resulting variants (Magi et al 2015). It largely deals with human datasets and processes pre-aligned short reads mapped to a reference genome to produce a very accurate set of SNP and indel calls for downstream analysis. The GATK pipeline

has two functionalities: a data processing pipeline and a variant detection pipeline. The data processing component includes recalibration of base quality scores, fixing miss-mapped reads proximal to or spanning indels, while the variant calling component generates a list of SNPs, indels and structural variations (SVs) by performing a local de-novo assembly. Genotype refinement is a very powerful GATK functionality, which assesses called variants based on concordance with data in publically available mutation/SNP databases such as dbSNP and the 1000 Genomes Project (McKenna et al. 2010; Campbell et al. 2012).

2.12.5. Variant annotation and filtering

A WES experiment generates a large amount of genetic variants many of which, in the context of a disease study, are not relevant as they are very unlikely to have a functional effect at the protein and/or systemic level. Methods for identifying the functional variants that have the potential to be involved in disease development or treatment response have therefore become very important, especially since use of WES is becoming increasingly prominent in personalized medicine applications. Functional SNP prediction tools such as ANNOVAR (Wang et al. 2010), dbNSFP (Liu et al. 2011) and SeattleSeq (http: //gvs.gs.washington.edu/SeattleSeqAnnotation/) are popular tools for annotating variants and predicting deleterious effects.

GeneTalk is a relatively new, professional network and online tool for analyzing human sequence variants (https://www.gene-talk.de/). It is used for annotation and filtering of variants and is based on publically available SNP/mutation databases such as dbSNP and NCBI.

ANNOVAR (http://www.openbioinformatics.org/annovar/) was specifically developed to annotate SNPs, indels and SVs from high-throughput sequencing data for classification of functional variation to generate a small subset of potential causative mutations resulting in protein-coding changes (Wang et al. 2010). Additionally, ANNOVAR annotates variants located in exonic, intronic, intergenic, splicing site, 5'/3' untranslated regions (UTRs), or upstream/downstream regions of genes and uses standardized nomenclature (Den Dunnen and Antonarakis 2000) to annotate non-synonymous SNVs and indels in relation to known cDNAs and proteins. Other important functions include measurements of evolutionary conservation calculated using PhastCons (Siepel et al. 2005) and PhyloP (Pollard et al. 2010), which can be implemented to predict the tolerability of a missense variant at a specific position. In addition to annotation of known and novel variants, ANNOVAR uses multiple algorithms to predict the functional effect of missense variants. Other such software tools include PolyPhen-2 (Ramensky et al. 2002), SIFT (Ng and Henikoff 2003), PROVEAN (Choi et al. 2012), LRT (Chun and Fay 2009) and MutationTaster (Schwarz et al. 2010). ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/) accessible

from the NCBI website, is a freely accessible, public archive of reports of the relationships among human variations and phenotypes, with supporting evidence from the literature.

SIFT serves the same purpose but is based on the degree of conservation of amino acid residues in sequence alignments derived from closely related sequences. Thus, by using sequence homology, it predicts the effects of all possible substitutions at each position in the protein sequence. SIFT is available as an online tool (http://sift-dna.org) and can be applied to naturally occurring non-synonymous polymorphisms or laboratory-induced missense mutations (Kumar et al. 2009). PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) predicts the possible impact of amino acid substitutions on the stability and function of human proteins using structural and comparative evolutionary considerations (Adzhubei et al. 2013). A more recent software tool called PROVEAN (Protein Variation Effect Analyzer), is documented to be comparable to SIFT or PolyPhen-2 (Choi et al. 2012). PROVEAN predicts whether an amino acid substitution or indel has an impact on the biological function of a protein and is useful for filtering gene variants to identify nonsynonymous or indel variants that are predicted to be functionally important (Choi and Chan 2015).

2.13. Genetic Counselling

Genetic counselling of patients considering genetic testing is essential to discuss the benefits and limitations of genetic testing. The value of breast cancer genetic testing does not only derive from the information obtained, but also from the ability to effectively communicate the implications of the test results to clinicians and their patients in ways that positively affect behavioural responses to receiving genetic information (Wilfond et al. 1997). Figure 2.6 illustrates the genetic counselling process implemented at Tygerberg Hospital, Western Cape, to ensure optimal use of scarce human resources. BRCA1/2 founder mutations are tested in a population-specific manner in South African breast cancer patients following consultation by the treating clinician (pre-test counselling). Post-test counselling by a registered genetic counsellor is performed after the BRCA report is provided in order to determine the appropriateness of 1) extended mutation analysis of the entire BRCA1/2 genes in mutation-negative breast cancer patients and 2) family screening in mutation-positive cases.

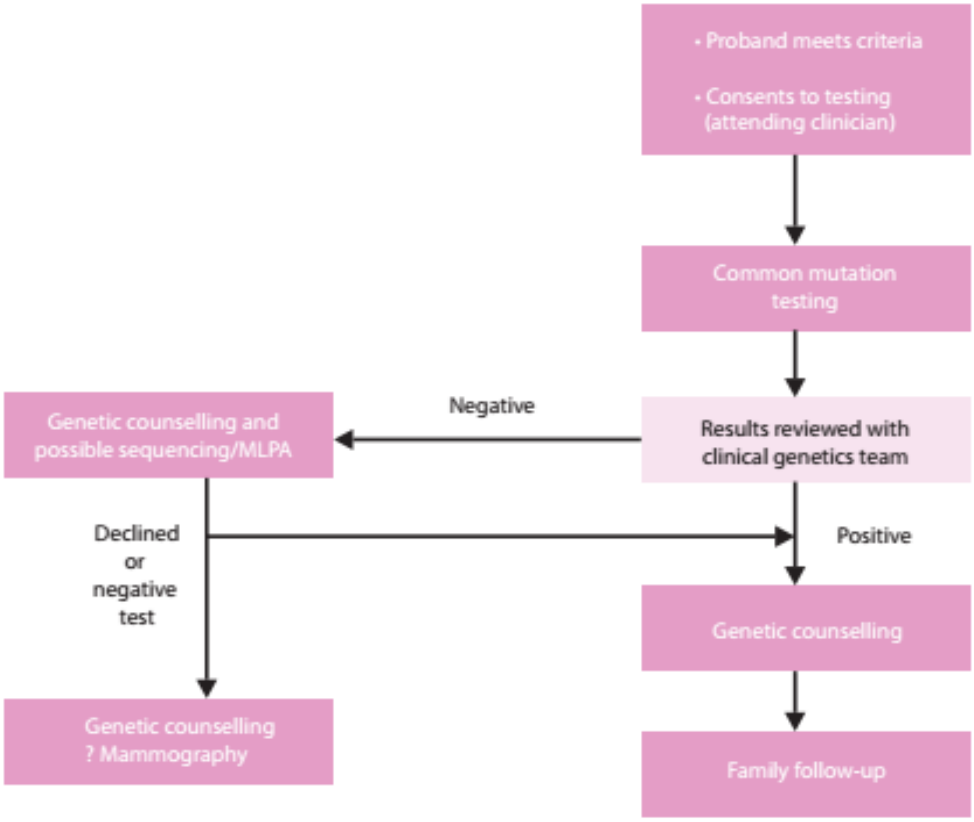


Figure 2.6. Genetic counselling and testing process in South Africa, 2010 – 2011 (reproduced with permission from Schoeman et al. 2013). Multiplex ligation-dependent probe amplification (MLPA) is a multiplex PCR method for the detection of abnormal copy numbers of genomic DNA or RNA sequences.

Since a positive DNA test may also imply increased risk in healthy family members, inadequate counselling may result in ethical, social, and legal implications. Genetic counselling should therefore form an integral part of the testing process in breast cancer patients (American Society of Clinical Oncology 1996). In South Africa, BRCA genetic testing is offered in accordance with guidelines developed in collaboration with genetic counsellors, surgeons and oncologists (Kotze et al. 2005; Schoeman et al. 2013), after careful consideration of potential ethical implications related to genetic discrimination (Kotze et al. 2004). Compliance to pre-determined guidelines and an ethics approved process for return of genetic information ensures that patients referred by clinicians are informed of the implications and limitations prior to the testing process (Kotze et al. 2015). When the test results are provided, appropriate medical management options need to be offered to the patient. With regard to pre-symptomatic genetic testing, a system should be in place preferably for ongoing monitoring of patients to measure the impact of treatment intervention in BRCA mutation carriers compared to sporadic or BRCA-negative familial cases. Managing patient expectations prior to genetic testing in relation to risk

implications and treatment options based on a positive or negative test result is of the utmost importance to ensure a meaningful genetic testing experience.

2.14. Risk assessment across the disease spectrum

Existing approaches to breast cancer prevention and treatment are being challenged by recent developments in genomic research which allows for a personalized approach to the prediction and management of disease across diagnostic boundaries. A genomics-based approach to breast cancer risk screening based solely on high-penetrance mutation screening is insufficient to explain familial inheritance related to breast cancer subtypes as well as adverse treatment-related side effects. Increasing recognition of the limitations imposed by a genomics-only approach to risk screening for breast cancer and associated comorbidities led to the development of a novel pathology-supported genetic testing approach (PSGT) which utilizes a combined research and service delivery platform to allow for the assessment of genetic risk in the context of relevant environmental determinants of phenotypic expression in a high-risk clinical setting (Kotze et al. 2015). Such a multidisciplinary approach to chronic disease risk screening has been proposed as a means of determining eligibility for extended genetic testing/WES performed alongside the evaluation of tumour morphology and hormone receptor expression profiles to allow for a subtype-specific approach to critical risk reduction intervention. Our PSGT approach is ideally suited as an intermediary step between GWAS and the clinical application of NGS used for the identification of novel causative variants of large clinical effect in genetically uncharacterized breast cancer patient subgroups.

One of the major challenges for translation of genetic research into clinical practice is the shortage of health practitioners educated in the field of genomics. A combined service and research approach that seeks to evaluate and learn from real world health outcomes of genomic applications was therefore applied in this study.

CHAPTER 3

AIMS AND RATIONALE OF THE STUDY

The present study was based on the hypothesis that the availability of a pathology-supported pharmacogenetics test for breast cancer, performed in conjunction with a medical and lifestyle assessment, may facilitate the diagnosis of treatable disease subtypes and prevention of cumulative risk with the ultimate aim to implement effective treatment strategies across diagnostic boundaries.

3.1. Aims

The study was performed in three phases aimed at the development of a comprehensive pathology supported genetic testing (PSGT) service for breast cancer informed by tumour pathology. The specific aims were to identify subgroups of breast cancer patients most likely to benefit from:

1) Pharmacogenetic CYP2D6 genotyping to reduce recurrence risk
2) Participation in a chronic disease screen to identify shared disease pathways
3) Exome sequencing to identify actionable mutations in genetically uncharacterised breast cancer patients

3.2. Rationale

Identification of a high-penetrance BRCA1 or BRCA2 mutation in female patients with breast cancer indicates a significantly increased risk for bilateral breast cancer and ovarian cancer. A positive test provides a mutation-specific assay for screening of close relatives to exclude or confirm the familial risk prior to implementation of preventive steps in affected individuals. Documentation of the age at diagnosis of breast cancer and evaluation of the family history are well-established indications of referral for BRCA mutation testing. A need however exists to develop a screening tool for selection of genetically uncharacterized breast cancer patients most likely to benefit from whole exome sequencing (WES).

The knowledge that breast cancer in women with germline BRCA1 mutations are mostly estrogen receptor (ER)-negative and typically lack expression of the progesterone receptor (PR) and overexpression of human epidermal growth factor receptor-2 (HER2) in their tumours, prompted the development of a pathology-supported pharmacogenetic test ideally suited to bridge the gap between single-gene and multi-gene testing using WES. This test performed as part of a newly-developed exome pre-screen algorithm (EPA) has the potential 1) to reduce recurrence risk caused by inappropriate/ineffective treatment due to variation in the CYP2D6

gene and 2) to prevent co-comorbidities in breast cancer patients frequently associated with cardiovascular risk and treatment-related side effects.

Given the clinical dilemma created by frequent detection of variants of uncertain significance (VUS) when using next generation sequencing (NGS), careful consideration is necessary before embarking on WES in any patient. Difficulties related to correct interpretation of the clinical meaning of VUS were addressed in this study by access to a genomics database resource developed at the interface between the laboratory and clinical practice. The research findings were published in scientific journals in the form of an original article and two congress abstracts listed below:

1. van der Merwe N, Bouwens CHS, Pienaar FM, van der Merwe L, Yako YY, Geiger DH, Kotze MJ (2012) CYP2D6 genotyping and use of antidepressants in breast cancer patients: test development for clinical application. Metabolic Brain Disease 27: 319-326

2. van der Merwe N, Pienaar R, van Rensburg SJ, Bezuidenhout J, Kotze MJ (2013) Identification of breast cancer patients at increased risk of “chemobrain”: Case study and review of the literature. S Afr J Psychiatry 19(3): 125

3. van der Merwe N, Peeters AV, van Rensburg SJ, Pienaar R, Bezuidenhout J, Kotze MJ (2015) Implications of discordance between reference genomes applied in breast cancer exome sequencing. Breast 24(Suppl 1): S121

The rationale for each of these sub studies is provided below in the form of an introduction to each phase of the study, in order to highlight the respective knowledge gaps addressed at the time. The focus of the two congress abstracts were modified during preparation of the full manuscripts prior to submission for publication in line with the information provided below.

3.3. PHASE I – Gene-drug Interaction

Depression is the most common psychiatric problem in cancer patients. As the age at diagnosis of breast cancer decreases in the general population, the risk of depression becomes higher (Kim et al. 2010). Younger patients suffer more from psychiatric distress, resulting in negative effects on quality of life and survival rate. The treatment of depression as part of a comprehensive risk reduction approach is therefore an important consideration in patients with breast cancer. It has been reported that between 20 % and 30 % of breast cancer patients treated with tamoxifen to reduce recurrence risk, also use antidepressants for depression or hot flashes (Nelson et al. 2006; Kim et al. 2010). Tamoxifen is the most commonly used anti-

estrogen drug for treatment of breast cancer, due to its active metabolites, hydroxytamoxifen and endoxifen. Selective serotonin reuptake inhibitor (SSRI) antidepressants can decrease the effectiveness of tamoxifen because these drugs compete for the cytochrome P450 2D6 (CYP2D6) enzyme that metabolises tamoxifen (Jin et al. 2005; Bernard et al. 2006). Consequently, antidepressants that act as CYP2D6 inhibitors may increase the risk of breast cancer relapse (Goetz et al. 2007). Antidepressants such as paroxetine, fluoxetine, and bupropion strongly inhibit CYP2D6 enzyme activity, whereas sertraline, duloxetine and diphenhydramine are considered moderate inhibitors. Escitalopram (Lexapro, Cipralex) venlafaxine (Effexor) and citalopram appear to have a very mild or no inhibiting effect on CYP2D6 activity (Preskorn et al. 2007; Lash et al. 2008; Holzman 2009).

Reduction in CYP2D6 activity due to enzyme inhibitors or polymorphisms in the CYP2D6 gene has been shown to reduce endoxifen levels (Borges et al. 2006) and may be associated with poorer outcome in women with breast cancer who have been treated with tamoxifen (Jin et al. 2005; Goetz et al. 2007; Newman et al. 2008). In addition to antidepressants and tamoxifen, CYP2D6 also metabolises antipsychotics, antiarrhythmic agents (propafenone, flecainide), beta-blockers (timolol, metoprolol, alprenolol), and opioids such as codeine and dextromethorphan. Due to CYP2D6 expression in the brain this enzyme furthermore mediates neuroprotection through inactivation of neurotoxins and break down of endogenous neural compounds such as catecholamines (Mann et al. 2011). Poor metabolisers have two copies of a defective CYP2D6 gene, and as a result they metabolise drugs at a much slower rate. This may result in a greater potential for adverse events and lower efficacy of drugs requiring CYP2D6 activation (Bernard et al. 2006).

Strong evidence has been provided that treatment failure due to impaired drug metabolism may increase the risk of recurrence of many different cancer types, not only in sporadic cases but also in patients with familial breast cancer (Newman et al. 2008). Various genes have been implicated in the development of breast cancer and those most studied include the BRCA1 and BRCA2 genes. Germline mutations in these two highly penetrant susceptibility genes explain the majority of familial breast cancer and additionally increase the risk for developing ovarian cancer (Sluiter and van Rensburg 2011). The lifetime risk for developing breast cancer is 65 % to 85 % for BRCA1 and 45 % to 85 % for BRCA2 mutation carriers (Chen and Parmigiani 2007). Increased frequencies of specific BRCA1 and BRCA2 mutations have been described in the South African Caucasian (Afrikaans-speaking) and Coloured (Mixed ancestry) populations (Reeves et al. 2004; Agenbag 2005; van der Merwe et al. 2012b). Three founder mutations [(BRCA1 c.1374delC (1493delC), BRCA1 c.2641G>T (2760 G>T, E881X) and BRCA2 c.7934delG (8162delG)] account for approximately 90 % of all BRCA mutation-positive families

in the Afrikaner population of European descent (van der Merwe and van Rensburg 2009). A relatively frequent BRCA2 founder mutation c.5771_5774delTTCA (5999del4) has also recently been identified in 3.4 % of Coloured and 25 % of Black Xhosa breast cancer patients studied in the Western Cape region of South Africa (van der Merwe et al. 2012b).

The vast majority of breast cancers in postmenopausal women are estrogen-receptor (ER)-positive (Punglia et al. 2008). However, in premenopausal breast cancer patients who are more likely to be carriers of BRCA1 or BRCA2 familial mutations, tumours display differences with regard to hormone receptors. While BRCA1 tumours are frequently ER and progesterone receptor negative, BRCA2 tumours are more commonly positive for both receptors (Lomen et al. 1998). The lack of hormone receptor positivity in most BRCA1 tumours suggests that treatment with tamoxifen or other hormonal therapy may be less effective in this patient population. In a study performed by Newman et al. (2008), in which tamoxifen-treated familial breast cancer patients were investigated for the effect of reduced CYP2D6 activity on clinical outcome, it was shown that the poor metabolizer status predicted worse overall survival in patients with familial breast cancer. Contrary to several other studies, the effect on enzyme function caused by both genetic variation in the CYP2D6 gene and concomitant use of a potent CYP2D6 inhibitor was taken into account. Patients with BRCA2 mutations had a significantly worse overall survival compared to patients with BRCA1 tumours (median survival: 7 versus 28 years; P=0.008; adjusted hazard ratio, 9.7). Based on the evidence provided in this study, it is of utmost importance to identify genetic subgroups of breast cancer patients most likely to benefit from CYP2D6 genotyping. Similar to BRCA mutation testing performed only in selected patients who fulfill the test criteria, CYP2D6 genotyping may only be applicable in a subgroup of breast cancer patients (e.g. patients with BRCA2- and ER-positive tumours considering concomitant use of tamoxifen and antidepressants).

In light of the increased frequency of particularly BRCA2 mutations in three different South African population groups and the fact that use of certain antidepressants can significantly reduce the function of the CYP2D6 gene, implementation of a pharmacogenetics CYP2D6 assay is an important consideration. The aim of this study (Phase I) was therefore twofold: 1) to determine the frequency of the relatively common inactivating CYP2D6*4 allele in South African breast cancer patients (validation data set) and 2) to determine the prevalence of depression/use of anti-depressants in breast cancer patients referred for genetic testing (implementation data set). Ultimately, breast cancer patients may benefit from the implementation of pathology supported genetic testing (PSGT) that combines relevant clinical and laboratory information for clinical decision-making. Although evidence based guidelines should ideally be developed before implementation of genomic applications, Khoury et al.

(2007) acknowledged the need to fill the remaining information gaps through ongoing data collection and health outcome studies. PSGT involves a combined service and research approach that seeks to evaluate “real world” health outcomes of genomic applications (Kotze and van Rensburg 2012), whereby overlapping aspects of translation research could provide feedback loops to allow integration of new genetic knowledge into clinical care (Khoury et al. 2007).

3.4. PHASE II – Gene-environment Interaction

Many studies to date have failed to demonstrate a significant overall survival difference between sporadic and familial breast cancer caused by mutations in the two major tumour suppressor genes, BRCA1 and BRCA2 (Bordeleau et al. 2010). These findings confirm the importance of lifestyle factors and genetic risk modifiers as determinants of adverse clinical outcome across breast cancer subtypes. Li et al. (2009) reported a more than seven-fold increased risk of a second primary contralateral breast cancer in women who smoke and consume seven or more alcoholic beverages per week. These authors found a 40-50% elevation in risk associated with obesity, which is consistent with the results from women enrolled in the National Surgical Adjuvant Breast and Bowel B-14 randomised trial of adjuvant tamoxifen performed in node-negative estrogen receptor (ER)-positive patients (Dignam et al. 2003). Breast cancer survivors in the highest quartile for body mass index (BMI) have at least 50% increased risk of a second primary breast cancer compared to women in the lowest BMI quartile (Trentham et al. 2007). King et al. (2003) demonstrated that the negative effect of obesity and lack of exercise also applies to breast cancer patients with familial breast cancer caused by mutations in the BRCA1/2 genes.

In breast cancer survivors, ten-year risk for cardiovascular disease (CVD) equals or exceeds that for cancer recurrence (Bardia et al. 2009). Genetic and environmental risk factors implicated in CVD furthermore overlap with those for major depressive disorder (Delport et al. 2014), multiple sclerosis (Davis et al. 2013) and Alzheimer’s disease (AD) (Luckhoff et al. 2015) as recently confirmed in the South African population, with the metabolic syndrome considered as a unifying risk factor (Kotze et al. 2015). This constellation of vascular risk traits including central obesity, hypertension, insulin resistance and dyslipidemia is also associated with different breast cancer subtypes (Rosato et al. 2011).

In view of the above, and given the genetic link found between BRCA1 mutations and heart failure (Shukla et al. 2011), it is clear that a simplified approach based purely on genetic information obtained from BRCA1/2 gene screening is not sufficient for risk management of familial breast cancer and other forms of lifestyle-related cancers. A pathology-supported

genetic testing (PSGT) strategy has proven useful for the evaluation of functional polymorphic variants as non-deterministic modifiers of disease risk and severity in patients predisposed towards CVD and AD (Kotze and van Rensburg 2012). In this context, a number of functional polymorphisms implicated in drug, lipid, iron, folate and homocysteine metabolism as well as thrombophilia were included in a CVD multi-gene assay found to be applicable in a wide spectrum of chronic diseases with a genetic component (Kotze et al. 2015). Integration of genetic information with relevant clinical, lifestyle and environmental factors as recently suggested by van der Merwe et al. (2012a) for breast cancer patients at increased risk of depression may therefore allow the development of a similar PSGT approach. Early detection of low-penetrance mutations or functional single nucleotide polymorphisms (SNPs) that affect the activity or amino acid sequence of BRCA modifier genes may facilitate the prevention of cumulative risk as a consequence of gene-environment interaction.

Several studies have shown that severe obesity is associated with increased risk for the development of hormone receptor negative breast cancer in pre-menopausal women (Petekkaya et al. 2013; Turkoz et al. 2013). In accordance with this notion, Kwan et al. (2015) also recently reported that obese patients have a higher likelihood of developing tumours associated with increased proliferation and a poor prognosis, including luminal B and basal-like breast cancer. Smoking and alcohol consumption were modestly associated with increased risk of ER-positive breast cancer (Kabat et al. 2011). Due to the known effects of obesity, alcohol and smoking on DNA methylation implicated in breast cancer and its associated comorbidities such as depression and CVD, genes involved in the folate-homocysteine pathway have been identified as important therapeutic targets in relation to both gene-diet (Heijman et al. 2003, Beilby et al. 2004) and gene-drug (De Mattia and Toffoli 2009) interaction. It is in this context of the emerging fields of nutrigenomics and pharmacogenomics that impaired folate homeostasis is considered a shared pathogenic mechanism for the risk of the metabolic syndrome, depression, vascular disease, cognitive impairment and breast cancer. In particular, a functional polymorphism (677C>T) in the methylenetetrahydrofolate (MTHFR) gene associated with disrupted one-carbon transfer and DNA mismatch repair is considered an important genetic risk factor for these medical conditions (Macis et al. 2007; Maruti et al. 2009; Peerbooms et al. 2011; Wu et al. 2013; Weiwei et al. 2014; Kakkoura et al. 2015). The low-penetrance MTHFR 677C>T mutation has also been linked to luminal B subtype of breast cancer (Naushad et al. 2012) in a similar way that BRCA1 mutations predominate in patients with the basal-like triple negative breast cancer subtype (Fostira et al. 2012; Rummel et al. 2013). MTHFR, methionine synthase (MTR) and methionine synthase reductase (MTRR) are among at least 30 different enzymes involved in the folate pathway. Altered enzyme activity may result in demethylation and increased cancer susceptibility (Choi and Mason 2002). Extensive alcohol consumption and

smoking are well-established triggers of the deleterious effects of variation in the MTHFR and apolipoprotein E (APOE) genes. While alcohol consumption in the presence of MTHFR gene variants have been documented as a risk factor for breast cancer, the APOE e-4 polymorphism has more recently been associated with breast cancer risk (Zunarelli et al. 2000; Moore et al. 2004; Chang et al. 2005). The APOE gene provides a genetic link between CVD and dementia, with the cholesterol-raising APOE e-4 allele considered an important risk factor for the metabolic syndrome, ischemic heart disease, stroke and AD. Together, APOE e-4 and MTHFR 677 C>T have been associated with treatment-related cognitive impairment in breast cancer patients.

Adherence to a Mediterranean-style diet high in folate-rich foods could increase levels of bioactive folate necessary for DNA synthesis, repair and methylation in MTHFR 677 T-allele carriers. Kakkoura et al. (2015) highlighted the biological importance of variation in the MTHFR and MTR genes in the association between the Mediterranean diet, the one-carbon metabolism pathway and breast cancer by acting as effect modifiers. In addition to its anti-neoplastic effects, folate may provide a host of benefits in patients with major depression including enhanced efficacy of antidepressants in non-responders (Papakostas et al. 2012). MTHFR 677 T-allele carriers also show a more optimal response to lifestyle-based interventions aimed at normalizing metabolic parameters implicated in increased vascular risk, including increased body weight and hypertension (Di Renzo et al. 2014). Individual genetic differences in the absorption and utilisation of such nutrients underscores the rationale for a PSGT approach that collectively incorporates the influence of multiple genetic and environmental risk factors in relation to protective nutrients in a new integrative medical model focused not only on cancer risk reduction, but health optimisation across the metabolic syndrome disease spectrum.

In this study, we sought to determine whether a PSGT strategy may be used to assess high BMI and other CVD risk factors in breast cancer patients stratified according to ER status, to facilitate the development of an exome pre-screening algorithm (EPA) for the selection of uncharacterized individuals set to derive optimal benefit from extended genetic testing including next-generation sequencing. A multidisciplinary approach to chronic disease risk screening incorporating a genomics component enabled consideration of both current and future risk implications in therapeutic decision-making across the diagnostic spectrum.

3.5. Phase III – From PSGT to Exome Sequencing

Microarray analysis of tumour genetics led to the identification of four intrinsic breast cancer subtypes, namely luminal A, luminal B, HER2-enriched and basal-like breast cancer (Perou et al. 2002) Next-generation sequencing (NGS) has however expanded the number of putative breast cancer subtypes from four to at least ten (Curtis et al. 2012; Ali et al. 2014). Given the

increasing availability and decreasing cost of NGS, this emerging technology is now able to provide clinicians and researchers with the opportunity to identify novel causative variants and potential drug targets across different breast cancer phenotypes (Patel et al. 2013). Whole genome (WGS) and whole exome sequencing (WES) can assist in the clinical diagnosis of breast cancer (Choi et al. 2009; Gahl et al. 2012), reveal the genetic basis of rare familial diseases (Ng et al. 2010; Bamshad et al. 2011) and allow for a novel approach to the characterization of disease pathology (Lupski et al. 2010). WES is an advanced method in the analysis of genes responsible for a subset of familial breast cancer (eg. P53 and CHEK2) that currently remains uncharacterized due to the laborious nature of a gene-by-gene sequencing process when mutations in the BRCA1/2 genes are excluded. In BRCA mutation-negative patients where early-onset breast cancer and drug failure or side effects cannot be explained by known mutations or functional single nucleotide polymorphisms (SNPs) expressed in a high-risk environment, WES performed in patients with familial breast cancer is a viable option for identification of potential novel causative genes/mutations.

The clinical utility of single-gene testing for causative mutations such as BRCA1/2 in a high-risk population group is well-established. However, the evaluation of multiple low-penetrance mutations/SNPs identified by genome wide association studies (GWAS) is complicated by uncertainty related to selection criteria and interpretation of the results for clinical application (Kotze et al. 2013). In this context, a pathology-supported genetic testing (PSGT) strategy was developed which allows for a multidisciplinary approach to chronic disease risk management used to inform clinical and therapeutic decision making across diagnostic boundaries (Kotze et al. 2015). This approach is facilitated by the assessment of genetic risk factors alongside relevant data concerning biochemical testing, tumour pathology and lifestyle factors. A combined research and service delivery approach provides an ideal platform for the assessment of current as well as future disease risk, based on the characterization of treatable chronic disease subtypes as an alternative approach to direct-to-consumer testing.

The clinical application of a validated multi-gene risk assay as part of this approach could also prove useful in the development of a pre-screen algorithm used to select genetically uncharacterized patients and/or those with severe treatment-related side effects eligible for extended genetic testing including WES (Kotze et al. 2015). This could allow for the identification of novel causative or risk-modifier genes in known pathogenic pathways in distinct high-risk patient subgroups across the disease spectrum.

The extension of the aforementioned PSGT platform to incorporate WES however presents a number of challenges related to technical issues associated with variegated aspects concerning the processing, analysis and interpretation of large-scale genetic data. These concerns highlight the pressing need to address challenges presented at every level of extended genetic testing including NGS to facilitate the integration of genomics-based chronic disease risk screening with WES. A pertinent issue which needs to be addressed in this context is the high level of discordance between different references genomes and variant-calling pipelines.

The commonly used National Centre for Biotechnology Information (NCBI) human reference genome hg19 (GRCh38) Pruitt et al. (2007) is derived from DNA samples obtained from anonymous donors and represents a small sampling of the vast array of human genetic variation. In addition, the detection of alternate alleles in high-throughput sequence data may be affected by preferential mapping of short reads containing the reference base over those containing an alternate base (Degner et al. 2009). Therefore, up to 30% of variants identified with hg19 are false-positives due to the presence of minor alleles at > 1.5 million loci, 4000 of which are associated with disease traits.

Dewey et al. (2011) developed three novel, ethnically concordant major allele reference sequences (MARS) from European [European ancestry from Utah (CEU)], African [(Yoruba from Nigeria (YRI)], and East Asian [(Han Chinese from Beijing and Japanese from Tokyo (CHB/JPT)] HapMap population groups using estimated allele frequency data at >7.9 million, >10.9 million and >6.2 million positions, respectively, catalogued in the 1000 genomes project. Low coverage pilot sequencing data comprising the majority of population-specific variation data had a sensitivity for an alternative allele of >99% at allele frequencies >10% and thus possesses high sensitivity for detecting the major allele (Durbin et al. 2010). Substitution of the ethnicity-specific major allele for the reference base resulted in single base substitutions at >1.5 million positions in all three populations. They demonstrated that the use of a MARS reduced genotyping error in common disease-associated variant loci by >40%.

The clinical application of WES in oncogenomic research and medical practice is vast and highly advantageous in the context of patient management. Discrepancies between different sequencing platforms and low concordance among multiple variant-calling pipelines however poses a major limitation which impedes the more widespread clinical adoption of extended genetic testing including WES (Boland et al. 2013; O' Rawe et al. 2013). In an attempt to overcome these obstacles we used the PSGT platform previously developed by Kotze et al. (2015) to formulate an exome pre-screen algorithm (EPA) for selection of genetically uncharacterised breast cancer patients set to derive optimal benefit from WES. After exclusion of high-impact mutations in the BRCA1 and BRCA2 genes in three patients selected for WES

using the EPA, we searched for potentially causative gene variants in known cancer-related genes that may match or contribute to the pre-determined clinical, pathology and genetic risk profile identified by PSGT using the newly developed EPA for breast cancer (Figure 5.1). The performance of the publicly-available hg19 and synthetic MARS was also compared between patients and controls.

CHAPTER 4

SUBJECTS AND METHODS

4.1. Ethical Approval

Ethical approval was granted for this translational research project by the Ethics Review committee of the University of Stellenbosch. Reference number N07/07/158 relates to recruitment of breast cancer patients at Tygerberg Hospital referred to as the validation data set. Reference number N09/08/244 relates to inclusion of participants in a chronic disease screening program referred to as the implementation data set. These were selected from the Gknowmix database generated at the interface between the laboratory and clinic using a combined service and research approach.

4.2. Study population

The study population consisted of **813 individuals** (Figure 4.1), including 175 South African breast cancer patients from two different population groups: 62 Coloured patients of Mixed Ancestry and 113 Caucasian patients of European Ancestry. Eighty eight cases (implementation data set) were selected from **625** South African Caucasian individuals who participated in a genomics-based chronic disease screening program (2010-2015) after providing informed consent for inclusion of their data in a secure centrally maintained research database (accessed at www.gknowmix.org). The remaining **87** patients (validation data set) were recruited at the Tygerberg Hospital Breast Cancer Clinic (Agenbag 2005). Of the 625 chronic disease screen participants, 214 controls without a family history of cancer were selected from the genomics database for comparative studies. A further **101** elderly females above the age of 65 years and without a diagnosis of cancer were included in the study.

Phase I of this study involved pharmacogenetic testing of 87 breast cancer patients from the validation data set and 23 patients from the implementation data set, as previously described by van der Merwe et al. (2012a). When this study was performed, the implementation data set merely consisted of 23 breast cancer patients which was later expanded as new samples became available. Agenbag performed full BRCA1/2 mutation screening in DNA samples from the validation data set, while assessment of human epidermal growth factor receptor-2 (HER2) status was not performed in these breast cancer patients at the time. Breast cancers were tumour biopsy-confirmed, and results on ER and progesterone (PR) status based on immunohistochemistry (IHC).

Phase II includes a subset of the validation data set (including a total of 87 patients) comprising 83 (60 Coloured and 23 Caucasian) breast cancer patients for whom ER-status was available. This patient group was combined with 81 (with known ER status) of the 88 breast cancer patients in the implementation data set derived from the chronic disease risk screening program. HER2 status (determined only in the implementation data set) was confirmed by

fluorescence in situ hybridisation (FISH) in equivocal IHC2+ cases. Age was used as a proxy for menopausal status using a cut-off age of 50 years to define postmenopausal status. All patients were screened using a CVD multi-gene assay first described by Kotze and Thiart (2003) to determine the extent to which treatment-related comorbidities matched or were discordant with the test results. A combined service delivery and research approach was used to identify breast cancer patients in the implementation data set considered for BRCA mutation screening based on well-established referral guidelines including early age of onset and a family history of cancer (Kotze et al. 2015).

Phase III of this study included the total number of 175 breast patients (87 from the validation data set and 88 from the implementation data set) as well as the total number of 315 control individuals (101 elderly females above the age of 65 years and 214 females participating in the chronic disease screening program for validation of variants identified by whole exome sequencing (WES).

Inclusion criteria: Patients histologically diagnosed with familial or sporadic breast carcinoma.

Exclusion criteria: Patients with benign neoplasms of the breast or other malignancies as well as those who did not provide consent for genetic testing.

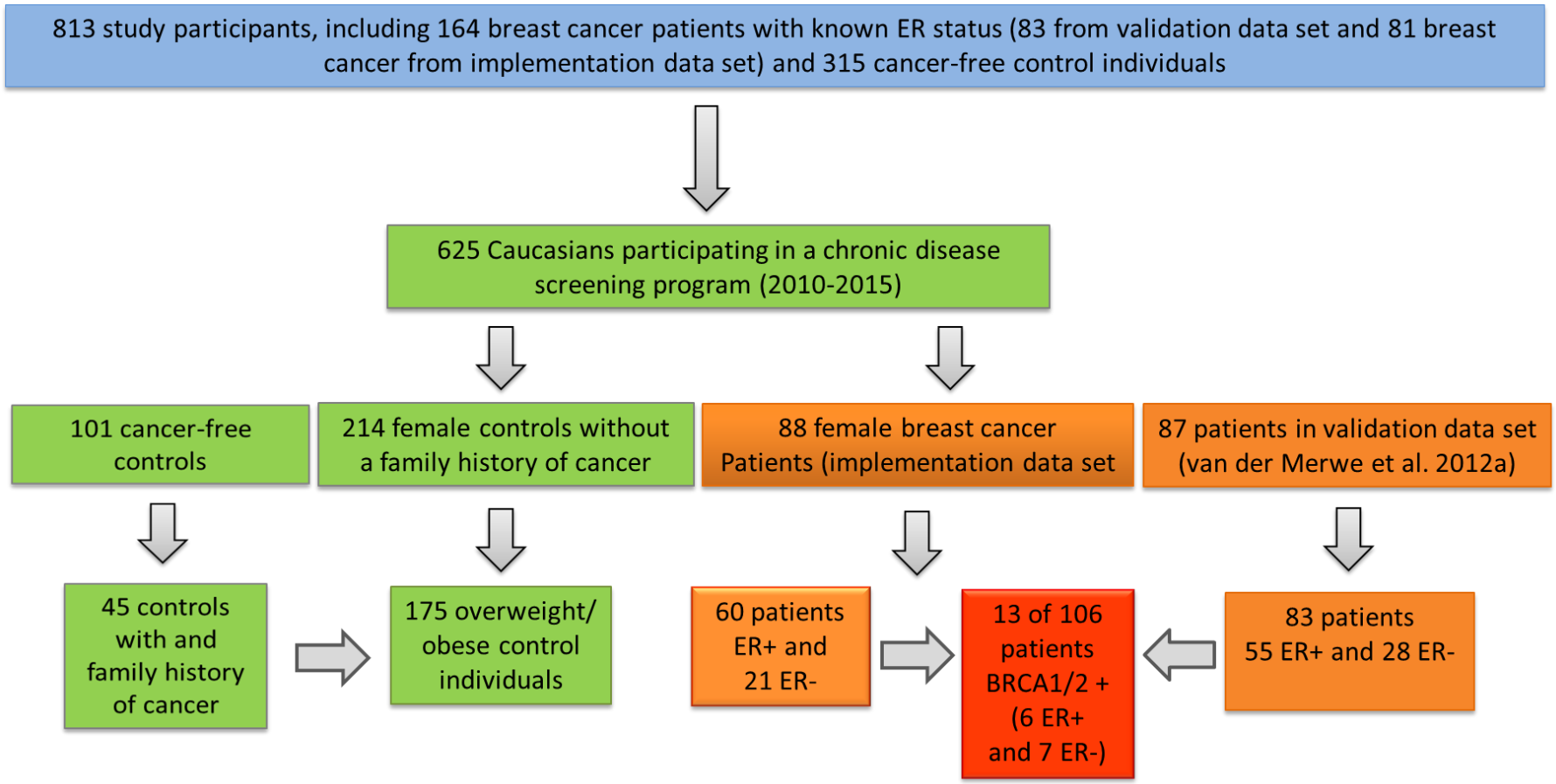


Figure 4.1. Description of the study population including patients participating in a chronic disease screening program and a validation data set described in van der Merwe et al. (2012a).

4.3. Questionnaire-based Nutrition and Lifestyle Assessment

All prospective participants of the chronic disease screen completed a medical history and lifestyle questionnaire (available at www.gknowmix.com) developed in collaboration with a registered dietician. This was used to document socio-demographic and clinical data including 1) age of breast cancer diagnosis/onset and associated comorbidities as well as, 2) disease recurrence, 3) family history of cancer, 4) tumour pathology and receptor status, 5) alcohol consumption, 6) body mass index (BMI), 7) smoking status, 8) level of physical activity, 9) and use of hormone replacement therapy (HRT), 10) oral contraceptive pills (OCPs) and medications such as antidepressants.

The questionnaire further denoted the number of days per week participants ate at least 5 fruits and vegetables and certain foods rich in fiber and folate over the course of the preceding three months. This information was used to calculate a fruit, vegetable and fiber score, which was considered to be low when less than 13, moderate between 14 and 18, and high when above 19. The folate score was considered to be low when less than 11, moderate between 11 to 13, and high when above 13. Intake of saturated/trans fat was considered low below 21, moderate from 22-26, and high above 27. An optimal BMI ranges between 18.5-24.9 kg/m^2 and individuals with a BMI above 30 kg/m^2 are generally classified as obese. Physical activity was documented for both day-time / occupational activity and participation in recreational sport (occasionally or at least 1 time per week), or structured exercise on a weekly basis (none, 2-3 times or 4 or more times) with each activity lasting more than 30 minutes. A physical activity score was then calculated from these two assessments.

The metabolic syndrome was defined based on the revised 2006 International Diabetes Federation (IDF) criteria (Alberti et al. 2006) (Table 4.1). Any additional information required was retrieved from patient records and included in the database. Written informed consent was obtained from all study participants.

Table 4.1 Clinical features for diagnosis of the metabolic syndrome based on three or more of five characteristics defined by universal evaluation values

Diagnosis of MetS Components	5 Clinical Indicators	Universal Evaluation Values	Personal Medical History (questionnaire)
Central Obesity	Waist circumference	>102 cm (Males) >88 cm (Females)	Obesity/overweight BMI > 30 kg/m^2

Diabetes type II	Glucose, fasting	≥5.6 mmol/L	Insulin resistance Type II diabetes
Dyslipidaemia	Triglycerides	≥1.7 mmol/L	Dyslipidaemia (excluding familial hypercholesterolaemia/high cholesterol)
	HDL-cholesterol	<1.0 mmol/L (Males) <1.3 mmol/L (Females)	
High Blood Pressure	Systolic	≥130 mmHg	Hypertension

4.1. DNA extraction and spectrophotometry

The DNA samples used in the study were generally extracted from whole blood using the QIAGEN QIAamp® DNA Blood Maxi Kit DNA extraction protocol (spin protocol) for whole blood (Appendix I). Subsequent to DNA extraction, the NanoDrop® ND-1000 spectrophotometer (v3.5.2 software package) (NanoDrop®, Technologies, USA) was used to determine the concentration and purity of the DNA samples.

Nuclease free water was used in the dilution of all the DNA samples to obtain a final concentration of 10 ng/µl, as required for Real-Time (RT)- PCR instruments.

4.5. Polymerase chain reaction amplification

Our initial focus was on CYP2D6 as the most widely studied pharmacogenetics marker of relevance in many clinical domains. Using polymerase chain reaction (PCR)-based methods, molecular analysis included detection of the allele 4 variant in the CYP2D6 gene (1846 G>A). DNA control samples were amplified using conventional PCR. Sequencing of these PCR products served as the gold standard and provided a basis for screening the larger population group. The results were then compared to validate real-time PCR results. Real-time PCR was carried out using the Applied Biosystems® TaqMan® Genotyping assay with the ABI 7900HT system as well as with the Corbett Rotorgene.

In phase 2 of this study, CYP2D6 genotyping was extended to the application of the CVD multi-gene assay described by Kotze and Thiart (2003) at the Pathology Research Facility (PRF).

4.5.1. CYP2DD6 allele 4 oligonucleotide primers

Oligonucleotide primers were designed to screen specific exonic regions of CYP2D6 allele 4. Reference sequences were obtained from the National Centre of Biotechnology (NCBI, www.ncbi.nlm.nih.gov) and/or UCSC Genome Browser (https://genome.ucsc.edu/). Using the BLAST (Basic Local Alignment Search Tool) function in NCBI, primers for CYP2D6 allele 4 (rs3892097) (Table 4.2) used in conventional PCR experiments were designed. Custom made primers and probes used for real-time PCR where designed and supplied with the Roche LightCycler® 480 II and ABI 7900HT genotyping assays (section 4.5.3).

Table 4.2. CYP2D6*4 primer parameters for conventional PCR experiments

Primer Name	Sequence (5' to 3')	GC Content (%)	T_M (°C)	T_A (°C)	Amplicon Size (bp)
CYP2D6*4_Forward	AGGGAGCAAGGTGGATGCAC	60.0	60.1	61	515
CYP2D6*4_Reverse	TAGGTCCAGCAGCCTGAG	61.1	56.4		

4.5.2. Conventional PCR Protocol and Cycling Parameters

The protocol used for analysis of CYP2D6 allele 4 is shown in Tables 4.3, with the cycling conditions for this assay shown in Table 4.4. DNA concentrations of samples varied therefore volumes of DNA added to the reaction mix were adjusted according to DNA content. Nuclease free water volumes were adjusted so that DNA and water amounted to a total volume of 18.9 µl.

Table 4.3. FastStart Fedility (Roche Diagnostics) PCR Protocol for CYP2D6*4 genotyping

Reagents	1x reaction (µl) for 50 µl reaction	1x reaction (µl) for 25 µl reaction	Final concentration
Nuclease-free water	36.8	18.4	
10x buffer without	5	2.5	1x
DMSO	2.5	1.25	5%
MgCl (25 mM)	2	1	1 mM
dNTP mix (10 mM)	1	0.5	0.2 mM
Forward primer (20 µM)	0.6	0.3	0.24 µM
Reverse primer (20 µM)	0.6	0.3	0.24 µM

FastStart Taq	0.5	0.25	2.5 U
DNA	1	0.5	50ng/µl
Final volume	**50**	**25**	

Table 4.4. Cycling conditions for conventional PCR experiments

Program	Cycles	Temperature	Time
Initial Denaturation	1	95°C	2 min
Amplification:			
Denaturation	28	95°C	30
Annealing		54°C	30
Extension		72°C	30
Final extension	1	72°C	4 min.

4.5.3. Real-time PCR protocol and Cycling parameters

The ABI™ TaqMan® SNP Genotyping assay for CYP2D6 1846G>A (rs3892097; ID C_27102431_D0) was used with the ABI™ 7900HT instrument for the real-time PCR experiments. This assay consists of prioritized primers and two allele-specific TaqMan® MGB probes to differentiate between the wild-type form and the SNP of the gene under investigation. Each probe is labeled with a different fluorophore, which is a fluorescent molecule that acts as a reporter dye. The fluorophores used by the probes in the assay are FAMTM and VICTM. The pre-designed assays are suitable for end-point genotyping by allelic discrimination analysis.

FAM-labeled probes only bind to the alleles that contain an adenine nucleotide (A) in genomic nucleotide position 1846 of the CYP2D6 gene, corresponding to the SNP (allele 4). FAM has an excitation wavelength of 470 nm and is detected in the green channel by the fluorometer of the ABI 7900HT due to its emission wavelength of 510 nm. VIC-labeled probes exclusively bind to the alleles that contain a guanine nucleotide (G) in genomic nucleotide position 1846 of the CYP2D6 gene, corresponding to the wild-type allele. VIC has an excitation wavelength of 530 nm and is detected in the yellow channel by the fluorometer of the ABI7900HT due to its emission wavelength of 555 nm.

The 40X TaqMan assay was diluted to 20X in double distilled SABAX water. The components of a 20 µl reaction were: 20 ng/µl of template DNA (2.0 µl), 10 µl of TaqMan Universal PCR Master Mix (P/N 4304437), 20X TaqMan SNP Genotyping Assay (1.0 µl) and 7.0 µl SABAX double

distilled water (Table 4.5). The amplification was carried out according to the standard Life Technologies ABI™ TaqMan® SNP Genotyping assay protocols. This includes a thermal cycling program consisting of an initial hold step at 95℃ for 10 minutes, 40 cycles of denaturation at 92℃ for 15 seconds and annealing/extension at 60℃ for 1 minute (Table 4.6), followed by the allelic discrimination run. Measurements were acquired at the end of the annealing/extension step of the run.

Table 4.5. CYP2D6 allele 4 standard TaqMan ABI™ allelic discrimination PCR reaction mix

Reaction Components	Volume/well (5µl Vol rxn)	Volume/well (20 µl Vol rxn)	Final concentration
Nuclease free water	0.25	7	
TaqMan universal PCR Master mix	2.5	10	1x
20x Genotyping assay mix	0.25	1	1x
Genomic DNA	2	2	

Table 4.6. ABI™ thermal cycling conditions for CYP2D6 allele 4 genotyping

Parameters:	Initial denaturation	denaturation	anneal/extend
Cycles	1	40	
Hold	10 min	15 sec	1 min
Target	95	92	60

4.6. Gel Electrophoresis

PCR products were resolved on a 2% (w/v) agarose gel to separate amplified DNA fragments based on the size of amplicons. The gel mix consisted of 2g agarose in 100 ml 1xTBE [90 mM Tris-HCl, 90 mM boric acid and 2.2 mM EDTA (pH 8.0)]. To enable visualization of the PCR products, 0.0001% (v/v) ethidium bromide (EtBr) was added to the electrophoresis buffer (70 µl of EtBr in 700 ml 1xTBE). A total volume of 8 µl of amplification product was loaded onto a gel consisting of 3 µl Ficoll Orange G loading buffer [0.1% (w/v) Orange G, 20% (w/v) Ficoll, 10 mM EDTA at pH 7.0] and 5 µl of amplification product. To ascertain the amplification of the correct PCR product, a molecular size marker (Promega 100 bp DNA ladder) was loaded together with

the PCR products on the agarose gel. Electrophoresis of the PCR products was performed for 45 minutes at 100 volts in 1xTBE buffer. Visualization of PCR products was achieved (using a Life Technologies TFX-35M trans-illuminator- GIBRO BRL) by ultraviolet light transillumination.

4.7. DNA Sequencing

PCR products obtained for CYP2D6*4 and were sent to the Central DNA Sequencing facility of Stellenbosch University for automated sequencing as well as post-PCR clean up and the electropherograms analyzed using FinchTV Version 1.4.0 (developed by Geospiza Research Team). The application of the software was to view the nucleotide sequences of each gene and compare it directly to the reference sequence as obtained from the NCBI (National Centre for Biotechnology Information) database for detection of SNPs.

4.8. Extended genotyping using the CVD multi-gene assay

Genotyping of 8 SNPs forming part of a CVD multi-gene assay initially described by Kotze and Thiart (2003) (Table 4.7) was performed as part of a chronic disease screening program using the standard operating procedures (SOPs) developed in the Pathology Research Facility, Stellenbosch University. ABI™ TaqMan® SNP Genotyping assays with the standardized reaction mixes listed in Table 4.5 and cycling parameters denoted in Table 4.6 was used on the Corbett Rotor-Gene™ 6000 and Roche LightCycler® 480 II. The primers and probes used were described by Fisher (2011). The gene variants listed in Table 4.7 selected for genotyping was extended to include MTR 2756 A>G (rs1805087), MTRR 66 A>G (rs1081394).

Table 4.7. Summary of the low-penetrance mutations evaluated as part of the CVD multi-gene assay representing the genetic component of a chronic disease screening program.

Genetic variation	Biological pathway	Disease Association	References
MTHFR 677 C>T, A222V (rs1801133)	Homocysteine and Folate metabolism	Various cancers including breast cancer	Liu et al. 2012; Shen et al. 2001; Graziano et al. 2006; Siemianowiczet al. 2003
		CVD	Di Renzo et al. 2014; Wald et al. 2002; Kotze and van Rensburg 2012
		Neurocognitive diseases including Alzheimer's disease, dementia, multiple sclerosis, cognitive impairment	Lewis et al. 2006
MTHFR 1298 A>C (rs1801131)		DVT	Elhassan and Abdalla 2015; Andreassi et al. 2006; Den et al. 2005 et al. 2015
		Recurrent pregnancy loss	Wu et al. 2012
APOE 3937 T>C, allele E4 (rs429358)	Lipid and Lipoprotein metabolism	Breast Cancer	Saadat 2012
		CVD	van Rensburg et al. 2000, Lengacher et al. 2015
APOE 4075 C>T, allele E2 (rs7412)		Neurocognitive diseases including Alzheimer's disease, dementia, Parkinson's and cognitive impairment	Seshadri et al. 1995; Anttila et al. 2004; Genin et al. 2011
Factor V Leiden 1691 G>A (rs6025)	Haemostasis	CVD	Volzke et al. 2005; Keijzer et al. (2002); Reznikoff-Etiévan et al. 2001; Mierla et al. 2012
		Recurrent pregnancy loss	Reznikoff-Etiévan et al. 2001; Mierla et al. 2012
Factor II 20210 G>A (rs1799963)		Blood clotting disorders including DVT	Elhassan and Abdalla 2015; Andreassi et al. 2006; Den et al. 2005
HFE 845 G>A, C282Y (rs1800562)	Iron metabolism and oxidative stress	Haemochromatosis	Bartzokis et al. 2011; Gurrin et al. 2008;
		Various cancers including breast cancer	Osborne et al. 2010; Liu et al. 2013
		CVD, cardiomyopathy	Klipstein-Grobusch et al. 1999; Kiechl et al. 1997; Roest et al. 1999; Silva et al. 2010
HFE 187 C>G, H63D (rs1799945)		Arthritis	Cauza et al. 2005
		Alzheimer's disease	Pulliam et al. 2003; Connor and Lee 2006; Ding et al. 2009
		Diabetes	Moczulski et al. 2001; Fernandez-Real et al. 1999; Kwan et al. 1998

4.9. Whole exome sequencing

Genetic analysis of breast cancer patients was initiated with BRCA mutation screening which was extended to 1) CYP2D6 genotyping for prediction of drug response/recurrence risk, 2) the chronic disease screen using the CVD multi-gene assay to assess cardiometabolic risk in patients with ER-positive versus ER-negative tumours and 3) to select patients for whole exome sequencing (WES) based on mutation status and the questionnaire based assessment (Kotze et al. 2015). Where extended mutation analysis of the entire BRCA1 and 2 genes as well as the CYP2D6 gene were warranted in cases where the most common variants generally tested were unable to explain breast cancer or the occurrence of drug side effects, WES to identify possible novel causative genes/mutations was performed rather than sequencing genes individually. Sequencing of BRCA1 and BRCA2 alone are particularly costly and does not include other causative cancer-related genes (P53, CHECK2, PTEN etc.).

Selected genetically uncharacterised breast cancer patients were subjected to WES to identify potential intermediate risk variants or 'BRCA-like' mutations that confer similar risk and may be responsible for the development of breast cancer. This process may only be initiated subsequent to following an ethically approved process including genetic counseling due to the complex nature of possible unintended information that may be obtained. As much of the exome data would be unrelated to breast cancer, feedback of results will only be focused on breast cancer and associated co-morbidities caused by a combination of genetic and environmental factors.

4.9.1. Cases selected for WES

In order to identify potential causative variants in genetically uncharacterized patients diagnosed with breast cancer, a combination of pathology and genetic tests was used to select individuals for WES upon obtaining ethical approval from Stellenbosch University (Ethics number: N09/08/224).

WES was performed at the Central DNA Sequencing facility of Stellenbosch University in three South African Caucasian patients and three population-matched controls. The BRCA1/2 mutation-negative index patient was diagnosed with invasive ductal carcinoma of the Luminal B subtype at the early age of 29 years. Her mother was diagnosed with bilateral (both ductal and lobular) breast cancer of the Luminal A subtype at the age of 57 years. The third patient included in this phase of the study was unrelated to the index patient and was diagnosed with triple-negative breast cancer at the age of 43 years (Table 4.7).

Table 4.7. Histopathologic classification of the three breast cancer patients.

Case	ER status	PR status	HER2 status	Tumour subtype
Index	Positive	Positive	Positive	Luminal B
Index mother	Positive	Positive	Negative	Luminal A
Unrelated patient	Negative	Negative	Negative	Basal-like

Next generation sequencing methods produce a large number of gene variants that need to be prioritized based on their potential deleterious effects. As no single tool can predict all the possible outcomes of effects of gene variants, this study has utilized multiple software tools based on different algorithms and resources to produce a prioritized list of functional variants

4.9.2. Whole exome capture and sequencing

WES was performed by the Ion Proton™ System for NGS, using already-extracted DNA (100ng/µl) from three breast cancer patients and three controls. The library preparation protocol used was the Ion AmpliSeq™ Exome RDY Library Preparation (Publication Number MAN0010084). Template amplification was performed using the Ion PI™ Template OT2 200 Kit v3 (Publication Number MAN0009133). Semi-conductor sequencing on Ion Proton system was performed using the Ion PI™ Sequencing 200 Kit v3 (Publication Number MAN0009136) with the Ion PI™ Chip Kit v2. This method is designed to target all human exons. The sequencing run time on the Proton™ System is approximately 4 hours which enables the entire workflow from library construction to primary data analysis and result generation to be done in under 24 hours. The Torrent Suite™ proprietary software is the standard program through which this NGS apparatus operates and offers simple run setup using predesigned workflows and run monitoring in real-time. The analysis pipeline generated FASTQ files as a quality control validation, reporting identified errors that occurred during data generation.

4.9.3. WES analysis

WES analysis comprised several intermediate steps involved in the read-processing pipeline (Figure 4.2). First, sequence reads were aligned to the publically available human reference genomes hg19 using the torrent mapping alignment (TMap) program. TMAP is alignment software for short and long nucleotide sequences produced by NGS technologies

(https://github.com/iontorrent/TS/tree/master/Analysis/TMAP). This was followed by read analysis and processing using the preinstalled Torrent Suite™ Software and variant calling with the Torrent Variant Caller plug-in. This Torrent Variant Caller plug-in is a web-based application that combines analytical tools and integrated biological and chemical content to identify and prioritize potential causal variants based on published data. Mapped reads are processed in BAM [indexed sequence alignment/map format (SAM)] file format. Furthermore, the Torrent Variant Caller also performs functional annotation and filtering of variants outputted in variant call files (VCFs). The finding that hg19 contains minor alleles at >1.5 million loci led to repeated mapping of the reads to a synthetic Caucasian major allele reference sequence (CEU-MARS) following the same process. Resulting VCFs were then processed using GeneTalk (https://www.gene-talk.de/), a web-based tool for filtering and annotation of uploadable VCFs.

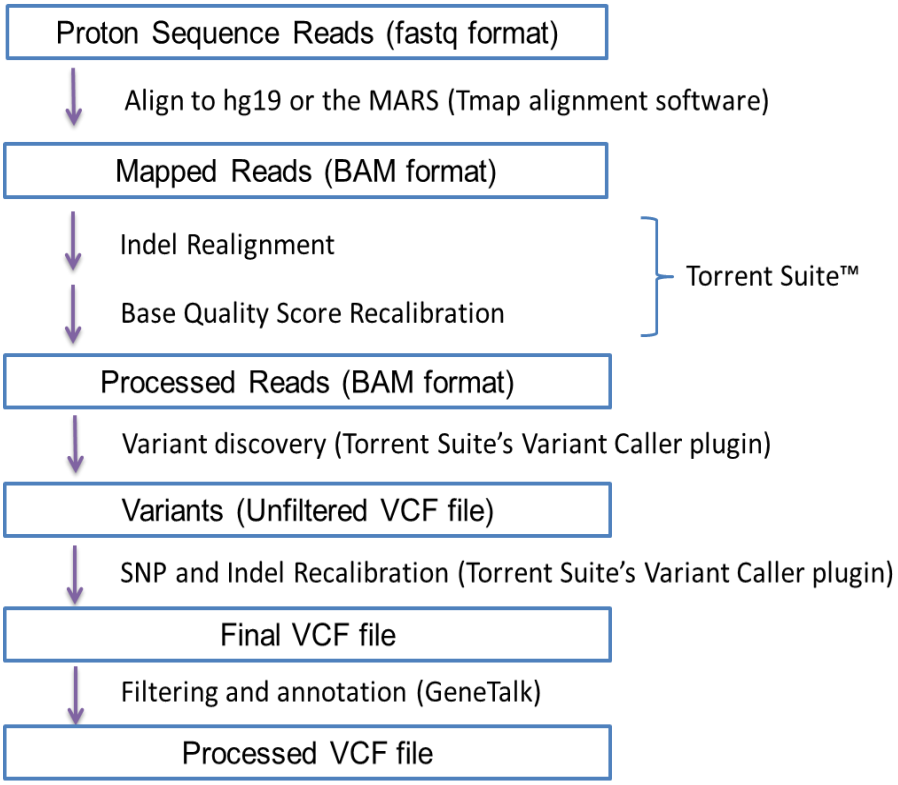
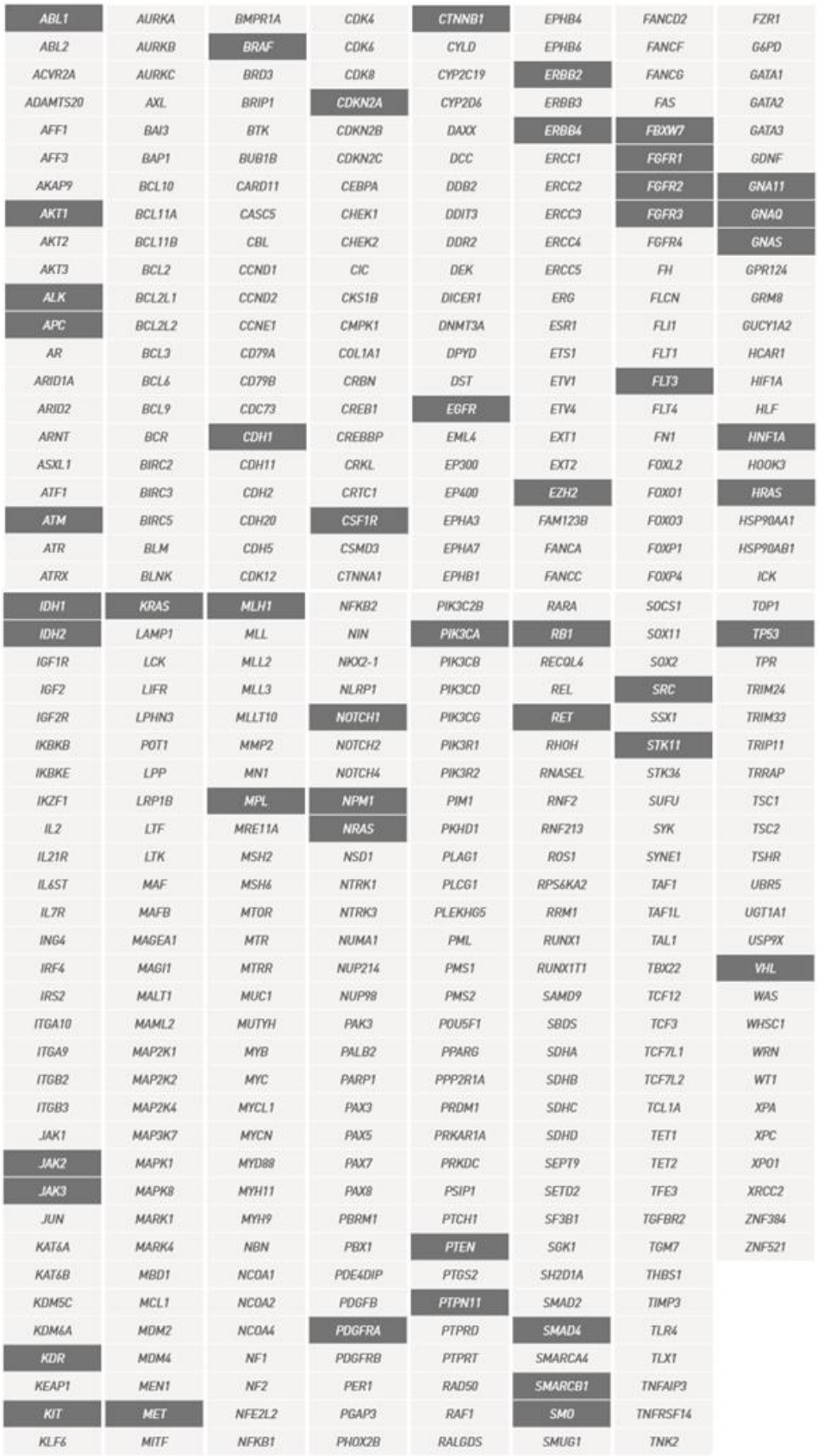


Figure 4.2. The next generation sequencing read-processing pipeline.

A minimum Phred quality score (Q score) of 50 was used for confidently calling bases, which translates into a base call accuracy of 99.999% (Ewing et al. 1988) - higher Q scores corresponds to higher quality. Furthermore, a minimum coverage depth of 100 was used for confident variant calls. Coverage provides counts of read depth (DP) at two different levels; the sample level where the DP value is the count of reads that passed the caller's internal quality

control metrics and at the site level where this value is the unfiltered depth over all samples. In addition to the aforementioned quality filters applied, variants were filtered on a population frequency of <0.01%, variant function (all synonymous variants excluded), and the most comprehensive cancer gene panel, the Ion AmpliSeq™ Comprehensive Cancer Panel including all-exon coverage of 409 genes (https://www.thermofisher.com/order/catalog/product/4477685) (Figure 4.3).



ABL1	AURKA	BMPR1A	CDK4	CTNNB1	EPHB4	FANCD2	FZR1
ABL2	AURKB	BRAF	CDK6	CYLD	EPHB6	FANCF	G6PD
ACVR2A	AURKC	BRD3	CDK8	CYP2C19	ERBB2	FANCG	GATA1
ADAMTS20	AXL	BRIP1	CDKN2A	CYP2D6	ERBB3	FAS	GATA2
AFF1	BAI3	BTK	CDKN2B	DAXX	ERBB4	FBXW7	GATA3
AFF3	BAP1	BUB1B	CDKN2C	DCC	ERCC1	FGFR1	GDNF
AKAP9	BCL10	CARD11	CEBPA	DDB2	ERCC2	FGFR2	GNA11
AKT1	BCL11A	CASC5	CHEK1	DDIT3	ERCC3	FGFR3	GNAQ
AKT2	BCL11B	CBL	CHEK2	DDR2	ERCC4	FGFR4	GNAS
AKT3	BCL2	CCND1	CIC	DEK	ERCC5	FH	GPR124
ALK	BCL2L1	CCND2	CKS1B	DICER1	ERG	FLCN	GRM8
APC	BCL2L2	CCNE1	CMPK1	DNMT3A	ESR1	FLI1	GUCY1A2
AR	BCL3	CD79A	COL1A1	DPYD	ETS1	FLT1	HCAR1
ARID1A	BCL6	CD79B	CRBN	DST	ETV1	FLT3	HIF1A
ARID2	BCL9	CDC73	CREB1	EGFR	ETV4	FLT4	HLF
ARNT	BCR	CDH1	CREBBP	EML4	EXT1	FN1	HNF1A
ASXL1	BIRC2	CDH11	CRKL	EP300	EXT2	FOXL2	HOOK3
ATF1	BIRC3	CDH2	CRTC1	EP400	EZH2	FOXO1	HRAS
ATM	BIRC5	CDH20	CSF1R	EPHA3	FAM123B	FOXO3	HSP90AA1
ATR	BLM	CDH5	CSMD3	EPHA7	FANCA	FOXP1	HSP90AB1
ATRX	BLNK	CDK12	CTNNA1	EPHB1	FANCC	FOXP4	ICK
IDH1	KRAS	MLH1	NFKB2	PIK3C2B	RARA	SOCS1	TOP1
IDH2	LAMP1	MLL	NIN	PIK3CA	RB1	SOX11	TP53
IGF1R	LCK	MLL2	NKX2-1	PIK3CB	RECQL4	SOX2	TPR
IGF2	LIFR	MLL3	NLRP1	PIK3CD	REL	SRC	TRIM24
IGF2R	LPHN3	MLLT10	NOTCH1	PIK3CG	RET	SSX1	TRIM33
IKBKB	POT1	MMP2	NOTCH2	PIK3R1	RHOH	STK11	TRIP11
IKBKE	LPP	MN1	NOTCH4	PIK3R2	RNASEL	STK36	TRRAP
IKZF1	LRP1B	MPL	NPM1	PIM1	RNF2	SUFU	TSC1
IL2	LTF	MRE11A	NRAS	PKHD1	RNF213	SYK	TSC2
IL21R	LTK	MSH2	NSD1	PLAG1	ROS1	SYNE1	TSHR
IL6ST	MAF	MSH6	NTRK1	PLCG1	RPS6KA2	TAF1	UBR5
IL7R	MAFB	MTOR	NTRK3	PLEKHG5	RRM1	TAF1L	UGT1A1
ING4	MAGEA1	MTR	NUMA1	PML	RUNX1	TAL1	USP9X
IRF4	MAGI1	MTRR	NUP214	PMS1	RUNX1T1	TBX22	VHL
IRS2	MALT1	MUC1	NUP98	PMS2	SAMD9	TCF12	WAS
ITGA10	MAML2	MUTYH	PAK3	POU5F1	SBDS	TCF3	WHSC1
ITGA9	MAP2K1	MYB	PALB2	PPARG	SDHA	TCF7L1	WRN
ITGB2	MAP2K2	MYC	PARP1	PPP2R1A	SDHB	TCF7L2	WT1
ITGB3	MAP2K4	MYCL1	PAX3	PRDM1	SDHC	TCL1A	XPA
JAK1	MAP3K7	MYCN	PAX5	PRKAR1A	SDHD	TET1	XPC
JAK2	MAPK1	MYD88	PAX7	PRKDC	SEPT9	TET2	XPO1
JAK3	MAPK8	MYH11	PAX8	PSIP1	SETD2	TFE3	XRCC2
JUN	MARK1	MYH9	PBRM1	PTCH1	SF3B1	TGFBR2	ZNF384
KAT6A	MARK4	NBN	PBX1	PTEN	SGK1	TGM7	ZNF521
KAT6B	MBD1	NCOA1	PDE4DIP	PTGS2	SH2D1A	THBS1	
KDM5C	MCL1	NCOA2	PDGFB	PTPN11	SMAD2	TIMP3	
KDM6A	MDM2	NCOA4	PDGFRA	PTPRD	SMAD4	TLR4	
KDR	MDM4	NF1	PDGFRB	PTPRT	SMARCA4	TLX1	
KEAP1	MEN1	NF2	PER1	RAD50	SMARCB1	TNFAIP3	
KIT	MET	NFE2L2	PGAP3	RAF1	SMO	TNFRSF14	
KLF6	MITF	NFKB1	PHOX2B	RALGDS	SMUG1	TNK2	

Figure 4.3. The Ion AmpliSeq™ Comprehensive Cancer Panel including all-exon coverage of 409 genes. The shaded blocks contain the 50 oncogenes and tumor suppressor genes that form part of the Ion AmpliSeq™Cancer Hotspot gene panel.

Resulting variants were subsequently evaluated in relation to publicly available mutation and SNP databases - including the Single Nucleotide Polymorphism Database (dbSNP) of the National Centre for Biotechnology Information (NCBI) (www.ncbi.nlm.nih.gov/SNP) and SNPedia (http://www.snpedia.com/index.php/SNPedia) that extracts information from PubMed on a daily basis to support genome annotation and interpretation - to verify their minor/risk allele status, prior to analytical validation of potentially deleterious gene variants in the laboratory. Clinvar on the NCBI (http://www.ncbi.nlm.nih.gov/clinvar/), Polymorphism Phenotyping v2 (PolyPhen-2), the Sorting Tolerant From Intolerant (SIFT) algorithm and the Protein Variation Effect Analyzer (PROVEAN) were used to predict the effect of variants on the biological function of resulting proteins. Finally, the obtained variants were confirmed with the use of the Integrative Genome Visualization (IGV) - a high-performance visualization tool used for interactive exploration of large, integrated data sets (www.broadinstitute.org/igv/).

4.10. Confirmation of WES-identified variants

Following the identification of potential causative gene variations with use of WES, confirmation was performed by the PRF laboratory by the PRF research team (Appendix 1). The selected variants were then screened for in the extended breast cancer patient (validation and implementation data sets) and control groups.

4.11. Statistical Analysis

Genotype distribution and allele frequencies were determined by allele counting and Hardy-Weinberg equilibrium (HWE) established using an exact test. Clinical and environmental risk factors relevant to the gene variants studied were compared between breast cancer patients from two different population groups, Caucasian and Coloured/Mixed Ancestry. Single and combined allelic effects of the selected SNPs were determined after adjustment for ethnicity in all cases, in addition to other potential confounding factors as appropriate. Possible alleleic combinations, together with their probabilities for each individual, were derived for selections of SNPs using functions from the R package haplo.stats. The qualitative characteristics were described using cross tabulation and frequency tables whereas the median and interquartile range were used for the quantitative phenotypes. Logistic regression models were used to compare pairs of groups, such as patients to controls, whereas linear regression models were used to compare the quantitative characteristics between the genotypes. Quantitative outcomes with non-symmetric distributions were log-transformations for analyses. The R software and R package genetics, freely available from http://www.r-project.org, were used for analyses. Results corresponding to p-values below 0.05 are described as significant.

CHAPTER 5

RESULTS AND DISCUSSION

The results of the study performed in three phases are presented below. In **Phase I**, we determined the appropriateness of combining CYP2D6 pharmacogenetic testing with diagnostic BRCA mutation screening aimed at the reduction of recurrence risk in patients with familial breast cancer. In **Phase II**, cardiometabolic risk factors were assessed in breast cancer patients stratified according to ER status, in order to facilitate the development of an exome pre-screening algorithm (EPA). The knowledge derived from these two studies was used in **Phase III** of this investigation to select three breast cancer patients from two unrelated families for whole exome sequencing (WES) in comparison with three cancer-free controls. We furthermore conducted a comparative effectiveness study using the standard human reference genome (hg19) and a synthetic major allele reference sequence (MARS), whilst investigating both rare variants in a familial context and low-penetrance single nucleotide polymorphisms (SNPs) at the population level as potential pharmacogenetic biomarkers that may be unique to each individual patient studied.

5.1. PHASE I RESULTS

The clinical characteristics of the 87 breast cancer patients previously screened for mutations in the BRCA1 and 2 genes are summarised in Table 5.1 The majority of patients were postmenopausal. Approximately half of these patients in the validation data set had a family history of cancer, with early onset (<50 years) breast cancer reported in 26% of patients.

Table 5.1. Clinical Characteristics of 87 Coloured and Caucasian female breast cancer patients included in the validation data set.

Parameter	Coloured patients (n=62)	Caucasian patients (n=25)	Total Group (n=87)
Mean age (yrs)	54	57	55
Family cancer history (yes)	27 (44%)	18 (72%)	45 (52%)
BRCA mutation positive	2 (3%)	3 (12%)	5 (6%)
Premenopausal	17 (27%)	6 (24%)	23 (26%)
Postmenopausal	45 (73%)	19 (76%)	64 (74%)

The age at diagnosis of breast cancer in the study population was significantly associated with a family history of cancer ($p<0.01$), independent of various possible confounders including ethnicity, body mass index, alcohol intake, smoking or CYP2D6*4 genotype (Table 5.2). A family history of cancer, expected to be closely related to BRCA mutation status, reduced the

expected average age at diagnosis/onset of breast cancer with 8 years. A similar effect on age of breast cancer diagnosis was not observed in relation to the presence or absence of the CYP2D6*4 allele.

Table 5.2. Analysis of clinical, lifestyle and genetic factors in Coloured and Caucasian study groups, in relation to age at diagnosis of breast cancer.

P-values for tests of association with age at diagnosis		**P-values**
Single	Confounders	Single
Ethnicity		0.3360
Genotypes	Ethnicity	0.8334
Alleles	Ethnicity	0.5830
Family history of cancer	Ethnicity	**0.0082**
Family history of cancer	Ethnicity, CYP2D6 genotype	**0.0097**
Body mass index (BMI)	Ethnicity	0.0582
Body mass index	Ethnicity, CYP2D6 genotype	0.0610
Body mass index	Ethnicity, Smoking	0.0948
Body mass index	Ethnicity, CYP2D6 genotype, alcohol	0.3713
Body mass index	Ethnicity, genotype, alcohol, smoking	0.4072

Successful amplification of the DNA fragments spanning CYP2D6 allele 4 (1846 G>A) was obtained in the PCR reactions carried out using DNA controls K1-K5**.** The PCR products obtained with the conventional primers designed for detection of CYP2D6 1846 G>A were visualized with ethidium bromide in a 2% agarose gel, presented in figure 5.1. The sizes of the amplicons produced were 298 bp, following PCR amplification using CYP2D6 1846 G>T primer sets. No non-specific amplification was observed for any of the DNA samples and distinct bands were visible, which therefore deemed all products suitable for further analysis. The PCR protocol was standardised based on these results and the same parameters were then used in the amplification of additional DNA samples.

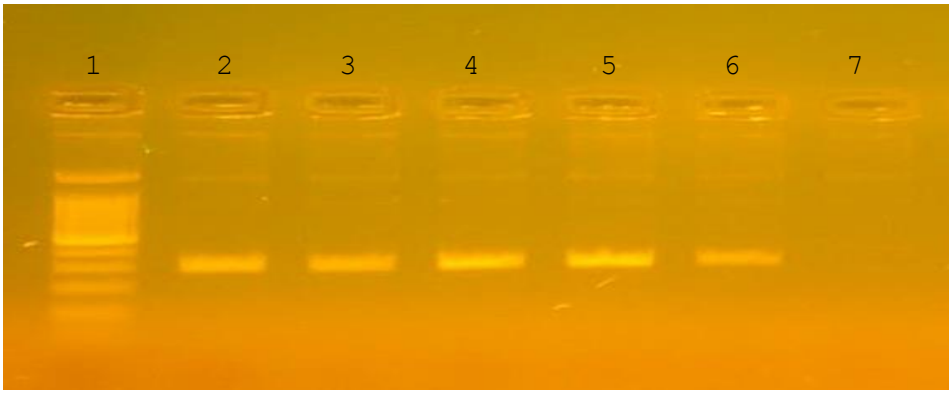


Figure 5.1: A 2% (w/v) agarose gel visualized with 0.0001% (v/v) ethidium bromide (EtBr). Lane 1 (from left to right) contains the DNA marker (100 bp) and while lanes 2-6 contain amplicons of 298 bp, following PCR amplification using the CYP2D6*4 allele (1846 G>T) primer set. Lane 7 contains the non-template control (NTC) PCR reaction product.

Figure 5.2 depicts the Sanger sequencing results for the CYP2D6*4 allele, performed by the Central Analytical Facility (CAF) of Stellenbosch University. The sequencing results are presented as electropherograms for the DNA controls. In the five DNA control samples (K1-K5) subjected to direct DNA sequencing, all three genotypes for the CYP2D6*4 allele were observed: three DNA samples were homozygous GG (figure 5.2), one heterozygous GA (figure 5.3) and one wild-type GG (figure 5.4). Both forward and reverse sequencing were performed, with the forward sequencing reactions depicted in the electropherograms. These control samples of known genotype were subsequently used to validate the real-time PCR methodology.

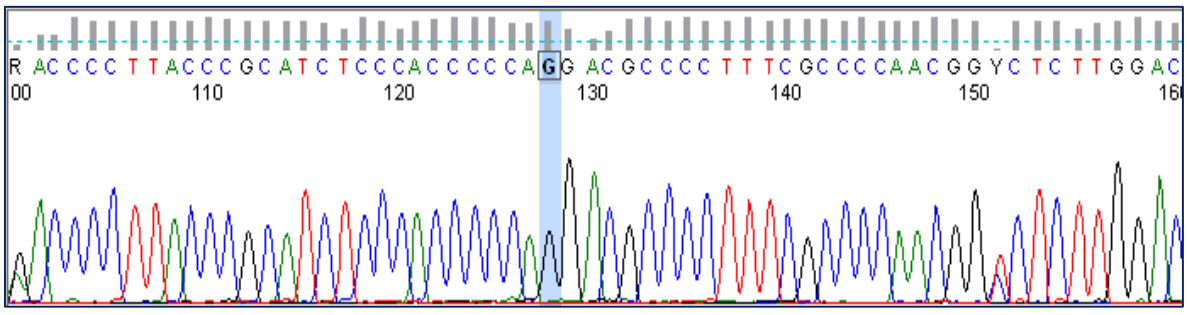


Figure 5.2. Electropherogram illustrating the forward sequencing reaction of an amplified PCR product of control sample K1 obtained with the CYP2D6*4 allele (1846 G>A) primer set. The

nucleotide position of the CYP2D6*4 allele is indicated by the highlighted region and corresponds to a genotype of GG, which represents a homozygous wild-type genotype.

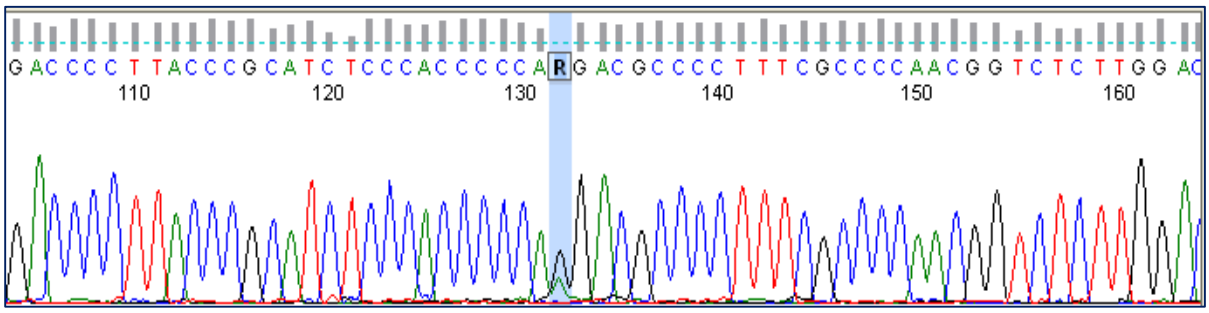


Figure 5.3: Electropherogram illustrating the forward sequencing reaction of an amplified PCR product of control sample K4 obtained with the CYP2D6*4 allele (1846 G>A) primer set. The nucleotide position of the CYP2D6*4 allele is indicated by the highlighted region and corresponds to a genotype of GA, which represents a heterozygous genotype.

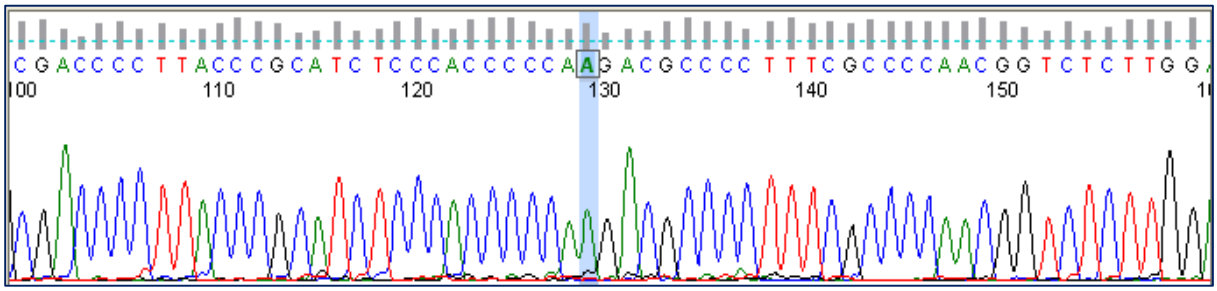


Figure 5.4: Electropherogram illustrating the forward sequencing reaction of an amplified PCR product of control sample K5 obtained with the CYP2D6*4 allele (1846 G>A) primer set. The nucleotide position of the CYP2D6*4 allele is indicated by the highlighted region and corresponds to a genotype of AA, which represents a homozygous mutant genotype.

Direct DNA sequencing confirmed the real-time PCR result of CYP2D6*4 homozygosity as obtained with the TaqMan genotyping assay using the ABI 7900HT, as well as the Corbett Rotor-Gene real-time PCR (Figures 5.5-5.8).

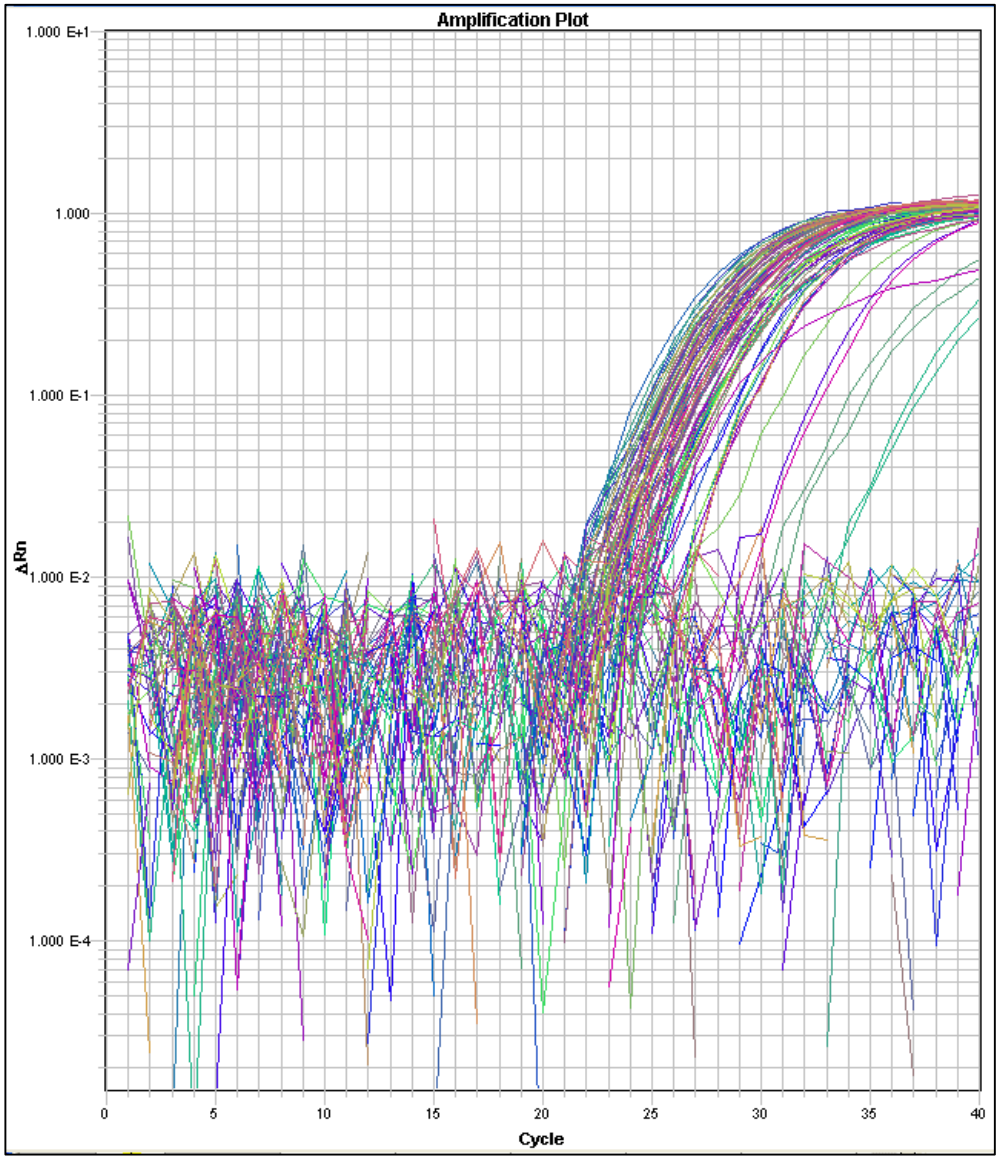


Figure 5.5. RT PCR amplification of sample batch one using the ABI *Taq*Man® CYP2D6*4 1834 G>A assay (ΔRn vs number of cycles). ΔRn = unit of fluorescence.

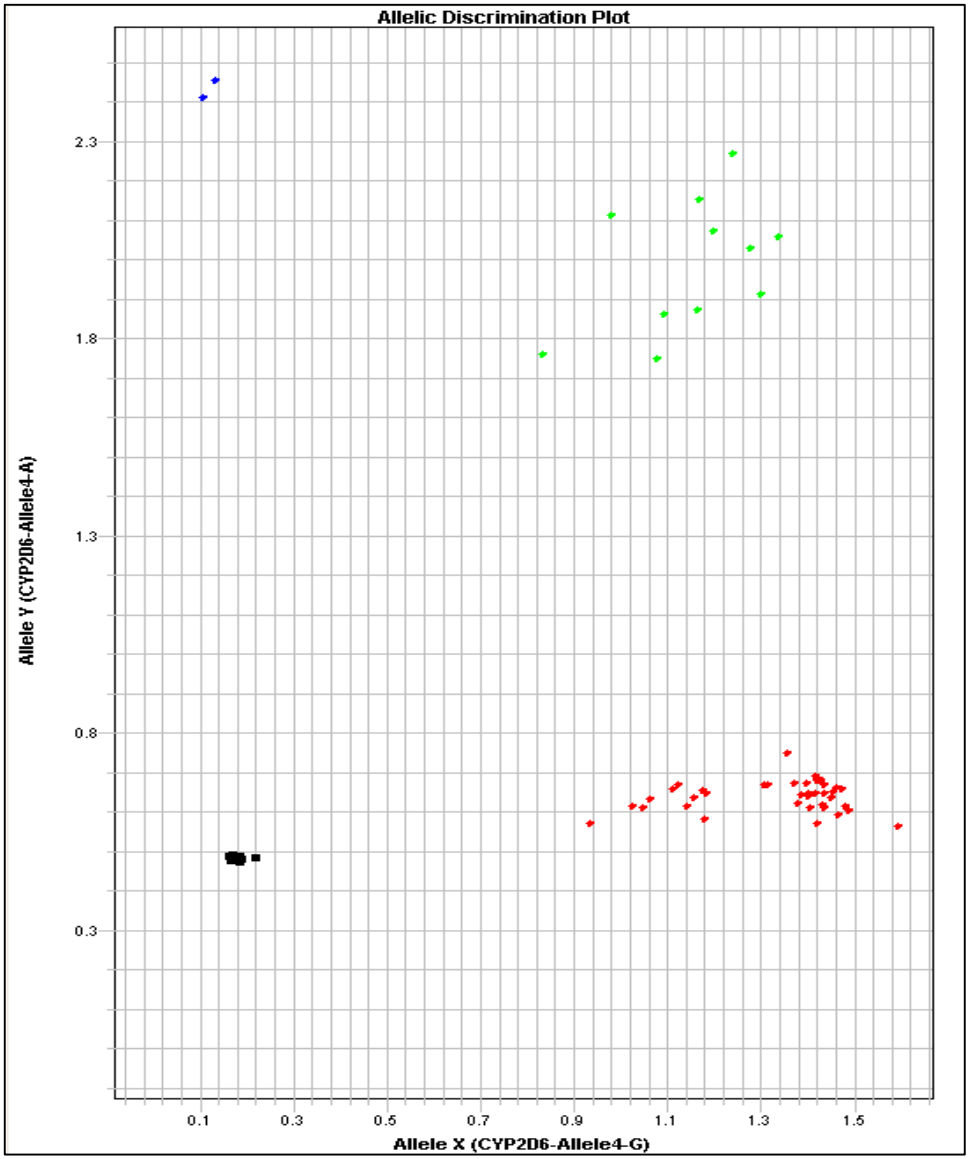


Figure 5.6. Allelic Discrimination Analysis for sample batch one using the ABI *Taq*Man® CYP2D6*4 1834 G>A [Allele Y (A- FAM labelled) vs Allele X (G- VIC labelled)]. Black = NTC, Red = Wild type, Green = Heterozygous, Blue = Homozygous

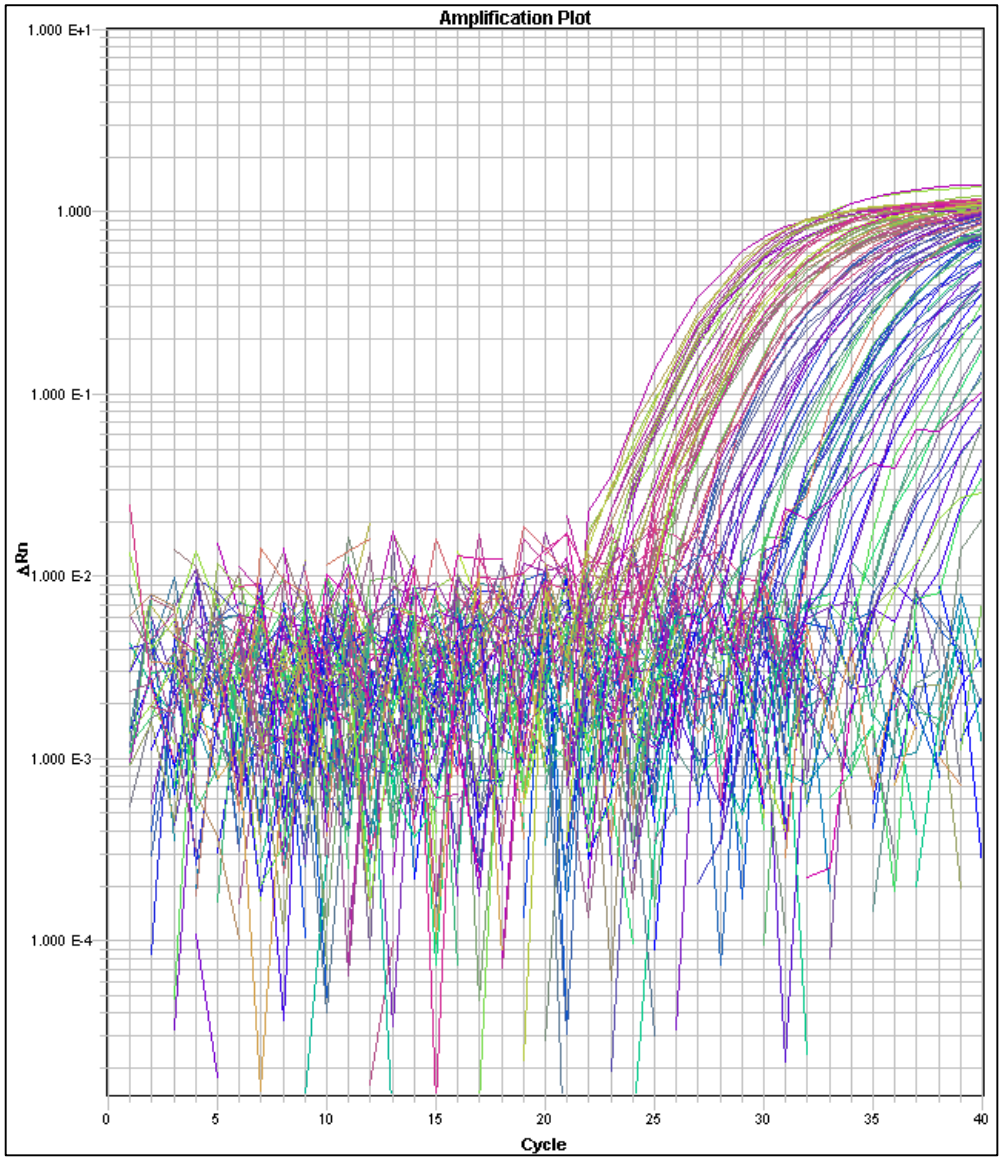


Figure 5.7. RT PCR amplification of sample batch two using the ABI *Taq*Man® CYP2D6*4 1834 G>A assay (ΔRn vs number of cycles). ΔRn = unit of fluorescence.

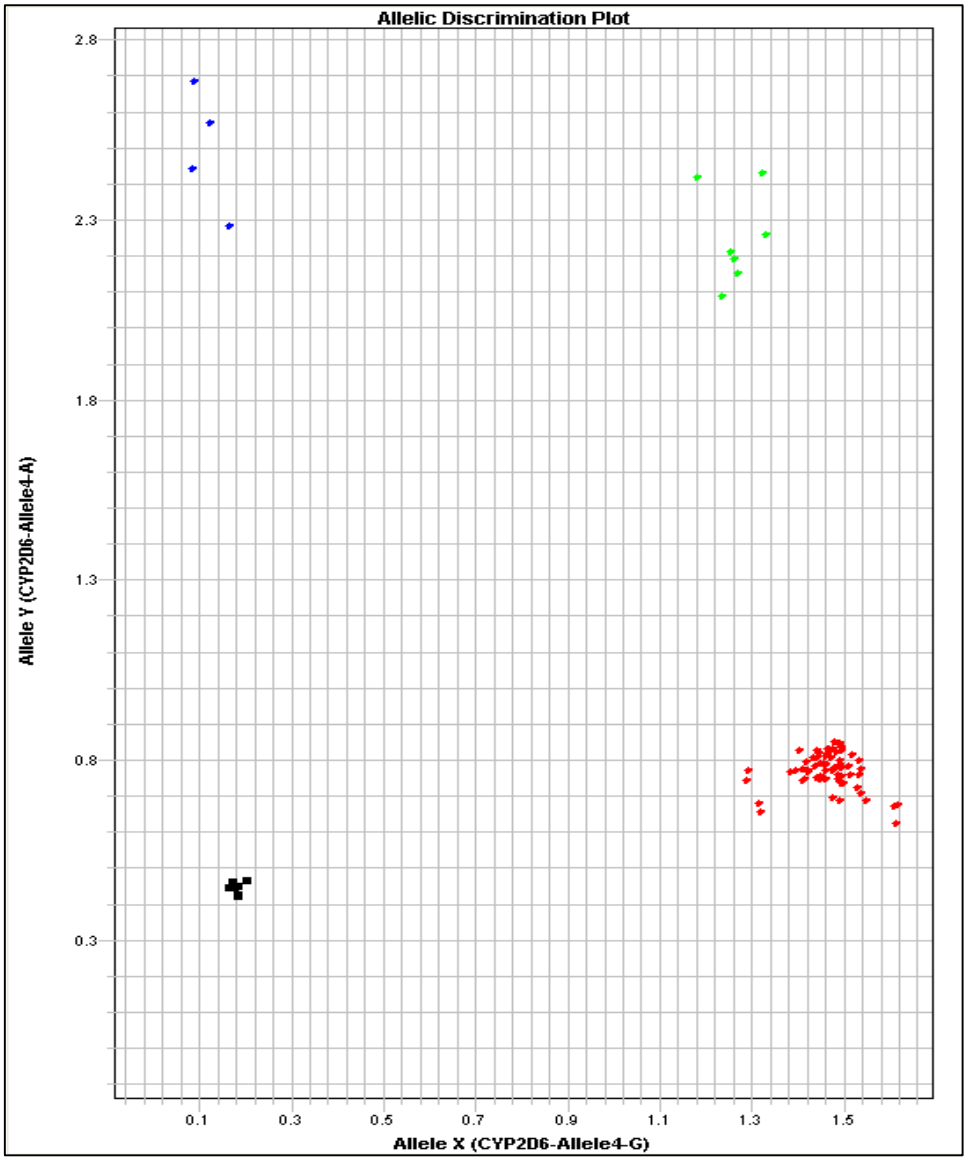


Figure 5.8. Allelic Discrimination Analysis for sample batch two using the ABI *Taq*Man® CYP2D6*4 1834 G>A [Allele Y (A- FAM labelled) vs Allele X (G- VIC labelled)]. Black = NTC, Red = Wild type, Green = Heterozygous, Blue = Homozygous

The allelic distribution of CYP2D6*4 was found to be in Hardy Weinberg Equilibrium in both the breast cancer patients and the controls. Genotype distribution and allele frequencies differed significantly between the Caucasian and Coloured population groups ($p<0.001$). The frequency of the CYP2D6*4 allele and the poor metaboliser status characterised by the homozygous genotype were highest in the Caucasian breast cancer patients. No significant differences in genotype distribution or allele frequencies were observed between the Caucasian patient and control groups (Table 5.3). The 21 male controls were excluded from this comparative analysis performed only in female breast cancer patients.

Table 5.3. Comparison of genotype distribution and allele frequencies for CYP2D6*4 between the Caucasian and Coloured study groups

	Control group (n=42)	Caucasian patient group (n=25)	Coloured patient group (n=62)	Caucasian patients to controls: P-value	Caucasian to Coloured patients: P-value
Allele					
G	70(83%)	38(76%)	120(97%)	P= 0.335	P< 0.0001
A	14(17%)	12(24%)	4(3%)		
Genotype					
GG	29(69%)	16(64%)	59 (95%)		
GA	12(29%)	6(24%)	2 (3%)	P= 0.282	P= 0.0001
AA	1(2%)	3(12%)	1(2%)		

As indicated in Table 5.4, two of the six (33 %) BRCA-positive patients included in the study were homozygous for the inactivating CYP2D6*4 allele. Both poor metabolisers (Caucasian) were ER-positive; one patient tested positive for a mutation in the BRCA1 gene and the other was BRCA2-positive. The BRCA2 mutation-positive patient (sample 22) provided an example of how CYP2D6 genotyping may affect clinical management. Instead of tamoxifen, an alternative endocrine treatment was administered after the patient underwent a hysterectomy due to her risk profile.

Table 5.4. Breast cancer patients with BRCA mutations in relation to CYP2D6 genotype.

Sample	Age	Ethnic group	ER status	BRCA mutation positive	CYP2D6*4
77	41	Caucasian	Negative	BRCA1	Wild-type
140	52	Caucasian	Positive	BRCA1	Homozygous
22	48	Caucasian	Positive	BRCA2	Homozygous
23	45	Coloured	NA	BRCA2	Wild-type
38	63	Caucasian	NA	BRCA2	Wild-type
66	63	Coloured	NA	BRCA2	Wild-type

NA, not available

The above-mentioned findings prompted extended CYP2D6 mutation analysis in an implementation data set of 27 breast cancer patients, to also assess the use of antidepressants in tamoxifen-treated breast cancer patients or patients who considered tamoxifen treatment. Ten patients (37 %) referred by participating clinicians reported a medical history of depression, with current use of antidepressants reported in four patients (15 %). None of these patients were homozygous for CYP2D6*4, while 4 heterozygotes were identified (Table 5.5).

Table 5.5. Breast cancer patients with a medical history of depression analysed during the implementation phase of the study.

Sample	Age	CYP2D6*4	Antidepressant
A	48	Wild-type	Not provided
B	58	Wild-type	Wellbutrin
C	47	Wild-type	Not provided
D	54	Wild-type	Zoloft, Wellbutrin
E	45	Wild-type	Cipralex
F	60	Heterozygous	Not provided
G	46	Wild-type	Not provided
H	59	Heterozygous	Not provided
I	68	Heterozygous	Not provided
J	57	Heterozygous	Cipramil, Cipralex, Wellbutrin

Knowledge of the different classes of CYP2D6 inhibitors is very important in the heterozygous patients as their ability to effectively metabolise tamoxifen could be affected by reduced enzyme activity caused by both the intermediate metaboliser status due to CYP2D6*4 heterozygosity and concomitant use of antidepressants. For example, one CYP2D6*4 heterozygous breast cancer patient (sample J) was diagnosed with depression at the age of 48 years and had been on various medications, including Wellbutrin (bupropion) known to strongly inhibit CYP2D6 function (Table 5.5). Drug side effects with use of antidepressants and various other medications were also reported. This ER-positive overweight patient with high cholesterol levels had bilateral breast cancer with 4 recurrent events between the ages of 39 and 57 years. She tested negative for three founder mutations in the BRCA1 and BRCA2 genes.

The pedigree of this ER-positive patient with two additional family members diagnosed with breast cancer is shown in Figure 5.9. She was referred by her oncologist for a comprehensive genetic screen that includes both a diagnostic and risk management component based on gene-diet (nutrigenetics) and gene-drug (pharmacogenetics) interactions, using a combined service and research approach. Further detail was obtained about the family history and age of cancer onset/diagnosis during the genetic counselling session (data not shown).

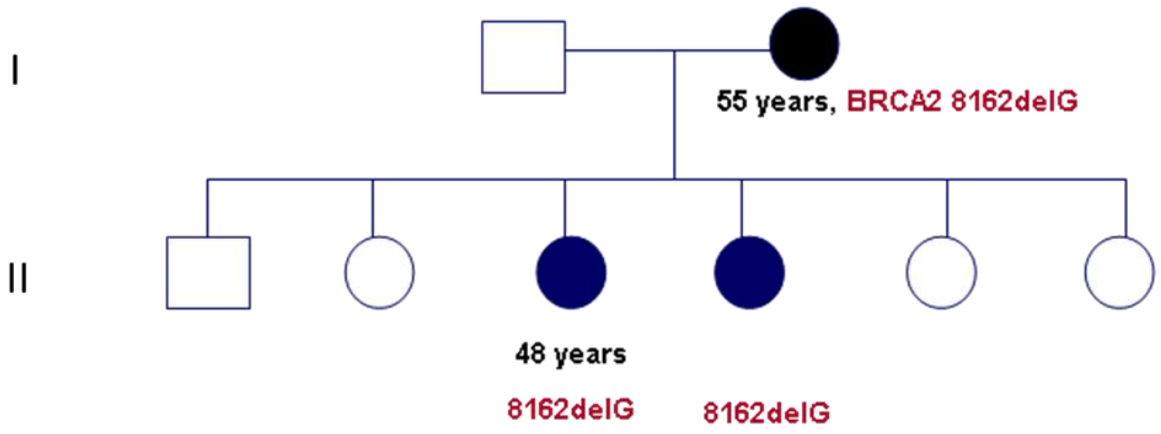


Figure 5.9. Pedigree of the index case diagnosed with breast cancer at the age of 48 years.

The genetic test was performed in conjunction with a medical and lifestyle assessment to identify a combination of risk factors that, if left untreated, could cause or contribute to disease development or recurrence. Table 5.6 provides a summary of the clinical characteristics and lifestyle risk factors entered into the Gknowmix Database at referral.

Table 5.6. Clinical and lifestyle information documented at referral of Sample 22

Family History		
Family Medical Conditions	Age of onset/diagnosis	Relationship
Stress / Anxiety / Depression		Sister
Breast cancer	55 years	Mother
Breast Cancer		Sister
Health Status		
Personal Medical Conditions	Age of Onset	Medication
Breast cancer, BRCA2 mutation positive	48 years	Considers Tamoxifen treatment
Clinical Assessment	**Value**	**Evaluation Values**
Weight	68.8 kg	
Height	1.73 m	
Body mass index (BMI)	22.99 kg/m^2	18.5-24.9 kg/m^2
Contraceptive pill	No	
Hormone replacement therapy	No	
Lifestyle Assessment	**Score**	
Physical activity	None / occasionally , Daily: Sedentary – Low	
Smoker	Current: No , Previous: Yes	
Alcohol consumption	1-13 Units / Week – Moderate	
Nutrition Assessment	**Score**	
Fat intake, saturated & trans fats	Moderate	
Fruit, vegetables, fiber intake	High	
Folate intake	Low	

5.2. PHASE I DISCUSSION

Identification of genetic subgroups of the population that are at increased risk of cancer initiation, progression or recurrence due to heritable susceptibility or ineffective therapeutic approaches is an important clinical consideration. Currently, DNA-based BRCA 1 and 2 gene testing and RNA-based transcriptional profiling including the determination of ER status are offered as a routine service to South African patients who fulfill the criteria for these tests (Kotze et al. 2005). This study aimed to determine the appropriateness of adding CYP2D6 genotyping to the breast cancer genetic test panel. The first step was to determine the frequency of the most common inactivating CYP2D6*4 allele in South African breast cancer patients (validation data set). The clinical usefulness of providing information related to CYP2D6 genotype to participating clinicians was furthermore assessed in a pilot study focused on the prevalence of depression and use of antidepressants in breast cancer patients subjected to diagnostic genetic testing (implementation data set). Since reduced CYP2D6 activity can be caused by variation in the CYP2D6 gene and/or the use of certain antidepressants, both aspects were considered in this study.

Breast cancer is most prevalent amongst South African Caucasian and Asian women and is the second most common cancer among Black and Coloured women (Vorobiof et al. 2001). This observation raises the question of whether genetic risk factors may explain ethnic differences in breast cancer risk. In this study, a significantly higher frequency of the CYP2D6*4 allele was detected in Caucasian (24%) compared with Coloured breast cancer patients (3 %). This finding is, however, unlikely to account for the higher prevalence of breast cancer in the Caucasian compared to Coloured group, since CYP2D6*4 affects drug response and has not been associated with an increased risk of primary breast cancer (Goetz et al. 2007). This is in accordance with the failure to detect a difference in allele frequencies for CYP2D6*4 between the Caucasian patient and control groups. Since CYP2D6*5 is relatively common in the Coloured population (Gaedigk and Coetsee 2008), future mutation analysis should be extended to include this assay. More than 75 CYP2D6 allelic variants have been identified to date (Bernard et al. 2006) and six inactivating alleles (*3, *4, *5, *6, *7 and *8) are responsible for approximately 95 % of the poor metaboliser status in most populations (Arneth et al. 2009).

The current debate in the literature as to whether CYP2D6 pharmacogenetic testing of breast cancer patients has clinical relevance highlights the importance of defining the selection criteria for a subgroup of patients where the potential benefits would outweigh the risk of unnecessary genetic testing. No statistically significant associations were observed between CYP2D6 genotype and recurrence risk in tamoxifen-treated patients in recent clinical trials (Rae et al. 2012; Regan et al. 2012). In contrast, strong evidence provided by numerous previous studies

showed that CYP2D6 genotype is directly (Newman et al. 2008) as well as indirectly (Jin et al. 2005) associated with the survival of breast cancer patients treated with tamoxifen (Kiyotani et al. 2010). In a study by Schroth et al. (2009) performed in 1,325 patients, recurrence rates were 14.9 % for extensive (wild-type genotype), 20.9 % for intermediate (heterozygotes) and 29 % for poor metabolisers (homozygotes) at 9 years of follow-up. All-cause mortality rates were 16.7 %, 18.0 % and 22.8 %, respectively. Conflicting results reported in the literature may be attributed to differences related to tamoxifen combination therapy, genotyping comprehensiveness, and CYP2D6 inhibitor coadministration (Hertz et al. 2012). These include frequent use of antidepressants prescribed to treat depression in breast cancer patients and/or menopausal vasomotor symptoms such as hot flashes. Tamoxifen treatment may induce or aggravate depression or anxiety, and commonly causes hot flashes (Kim et al. 2010). During the research-focused implementation phase of our study, it was found that 37% (10/27) of breast cancer patients from the implementation data set suffer from depression, while the use of antidepressants for depression were documented in four patients (15%). These findings are in agreement with previous studies which reported that between 20% and 30% of tamoxifen-treated breast cancer patients also use antidepressants (Nelson et al. 2006; Kim et al. 2010).

Tamoxifen is the standard 5 year endocrine treatment for premenopausal and postmenopausal women with ER-positive breast cancer. In contrast, aromatase inhibitors are frequently prescribed for postmenopausal women (Kim et al. 2010). The use of aromatase inhibitors has been prohibited in premenopausal breast cancer patients due to the activation of ovarian functions implicated in polycystic ovarian diseases. In premenopausal patients for whom the use of aromatase inhibitors is inappropriate, tamoxifen is the treatment of choice. Given the link between CYP2D6 genotype status and clinical outcome in risk allele carriers on tamoxifen, it is likely that aromatase inhibitors would be the preferred treatment for postmenopausal women with deficiencies in tamoxifen metabolism (Kim et al. 2010; Punglia et al. 2008). Newman et al. (2008) suggested that for postmenopausal patients with BRCA2 mutations and the CYP2D6 poor metaboliser genotype, an aromatase inhibitor would be the drug of choice.

The above-mentioned recommendation was of particular relevance in at least one patient with a founder BRCA2 mutation found to be homozygous for the CYP2D6*4 allele (sample 22). This ER-positive breast cancer patient reported a strong family history of breast cancer, which was shown in this study to be associated with an 8 year (on average) earlier onset of breast cancer irrespective of whether a mutation has been identified in the BRCA1 or BRCA2 genes. Based on this patient's CYP2D6*4 genotype and the fact that she had a hysterectomy, the treating oncologist prescribed an aromatase inhibitor instead of tamoxifen treatment. The proven clinical utility of CYP2D6 genotyping resulting in a change in clinical management in this patient, has

led to increased awareness of the potential consequences of CYP2D6 inactivation. This has also translated into avoidance of concomitant use of tamoxifen and inactivating antidepressants in an increasing number of breast cancer patients. The CYP2D6*4 heterozygote who had at least four episodes of recurrent breast cancer between the ages of 39 and 57 years (sample J) provides an example of where this information could have been relevant if provided when she first presented with bilateral breast cancer. She had been diagnosed with depression at the age of 48 years and has subsequently been treated with three different antidepressants, of which the use of Wellbutrin (bupropion) reported at entry into this study is known to strongly inhibit CYP2D6 activity. This ER-positive breast cancer patient reported side effects with use of antidepressants and various other medications. Arimidex (anastrozole), an aromatase inhibitor, was subsequently prescribed due to her current postmenopausal status. This case provides insight of how a patient's quality of life and well-being could be adversely affected by reduced CYP2D6 activity compromised further by use of antidepressants metabolised by CYP2D6. Extended mutation analysis of the CYP2D6 gene as well as the entire BRCA1 and 2 genes are considered in this patient to further investigate the cause of her poor treatment outcome.

The above-mentioned information explains what is meant by Khoury et al. (2007), who stated that evaluation of the interaction between providers and patients would enable the assessment of risks and benefits outside the context of randomized clinical trials. These authors acknowledged that while the translation research pathway for therapeutics is relatively straightforward, it is less clear for genetic tests that may provide useful information only in a subgroup of patients in some populations. Lorizio et al. (2011) found that CYP2D6 genotyping in breast cancer patients affects choice of therapy even in the absence of definitive data on clinical impact. This confirms the value of genetic information to guide treatment decisions on an individual basis, compared to a one-size-fits-all approach applied in most clinical trials. CYP2D6 genotyping is particularly relevant in the local population due to the increased frequency of BRCA2 founder mutations (van der Merwe et al. 2012a) and the high recurrence risk associated with reduced CYP2D6 activity in mutation-positive patients (Newman et al. 2008). Since CYP2D6*4 is virtually absent in the Black South African population (Wright et al. 2010), genotyping for this allele would not be appropriate in this particular ethnic group.

We believe that the same path used for development and implementation of genetic tests for high-penetrance genetic variations identified in the BRCA1 and 2 genes should be followed for low-penetrance genetic variations such as those in the CYP2D6 gene, with the exception that known environmental exposures must also be taken into account. Integration of CYP2D6 genotype information into the clinical decision-making process would require careful consideration of all possible determinants of CYP2D6 metabolic activity. Although all

determinants of CYP2D6 metabolic activity have not yet been identified and some are poorly understood, current knowledge related to the consequences of concomitant use of tamoxifen and CYP2D6 inhibitors is well-documented and needs to be communicated to clinicians and their patients. The phenotypic expression of genetic variation in the CYP2D6 gene is modulated by various external factors including physiological states (e.g. pregnancy) and certain herbal products. Borges et al. (2010) used genotype, concomitant medication and phenotype data to develop a CYP2D6 activity score that incorporates both CYP2D6 genotype and CYP2D6-mediated drug interactions. Correlation of this score with CYP2D6 phenotype represented a significant improvement over the use of CYP2D6 genotype alone.

Multiple gene/mutation testing in the context of the patient's medical and/or family history provides a basis for development and formulation of personalised cancer prognostication, drug treatment and lifestyle-related risk reduction programs (Cleator and Ashworth 2004). CYP2D6 genotyping has recently been included as the pharmacogenetics component of a multi-gene test for CVD (Kotze et al. 2003; Kotze and van Rensburg 2012), due to increased risk of muscle pain in patients taking cholesterol-lowering medication metabolised by CYP2D6 (Frudakis et al. 2007). Due to the potential significance of CYP2D6 genotyping in breast cancer patients and the increased risk of thrombophilia and cardiac complications with the use of chemotherapy and hormonal treatment (e.g. tamoxifen, HRT), the development and validation of a comprehensive screening strategy is hereby warranted to improve well-being and quality of life. In this study the PSGT approach - initially developed for risk management of CVD and related medical conditions (Kotze and van Rensburg 2012) - was applied during the implementation phase of this study to assess the clinical usefulness of combining diagnostic and pharmacogenetic testing in breast cancer patients. Assessment of CVD risk factors in breast cancer patients is supported by the direct link between an increased risk of postmenopausal breast cancer and the metabolic syndrome characterised by central obesity, insulin resistance, hypertension and/or hyperlipidemia (high triglycerides and/or low HDL-cholesterol levels) (Rosato et al. 2011). These findings support the implementation of CYP2D6 genotyping as part of a multi-gene test that takes co-morbidities such as obesity (King et al. 2003; Li et al. 2009), family history, ethnicity, ER status, BRCA mutation status and use of antidepressants into account prior to tamoxifen treatment of breast cancer patients. Although CYP2D6*4 is not deterministic by itself to cause disease, it may become clinically relevant in the presence of environmental co-factors or co-inheritance with other genetic risk factors (Mellick 2006; Newman et al. 2008).

We demonstrated the value of combining diagnostic BRCA mutation screening and pharmacogenetic testing in South African patients diagnosed with breast cancer. This study has contributed to an increased awareness of the important role clinicians can play in identifying

breast cancer patients who may benefit from CYP2D6 genotyping, based on their unique genetic background and concomitant use of tamoxifen and certain antidepressants. While it remains uncertain whether a pharmacogenetic profile should be obtained prior to initiating tamoxifen treatment, we conclude that genetic testing performed as part of a comprehensive risk reduction strategy could prove useful in informing clinical and therapeutic decision making in selected breast cancer patients. CYP2D6 genotyping has important relevance in several clinical domains and represents one of several pieces of information that clinicians may consider in guiding the selection of appropriate treatment tailored to the unique needs of the individual. This includes an important role in cardiovascular pharmacogenomics that may be relevant in some breast cancer patients. The weight of current scientific evidence in relation to risk-benefit assessment supports CYP2D6 genotyping in breast cancer patients who (1) are receiving tamoxifen and (2) are at high risk for tumour recurrence (e.g. family history, BRCA1/2 mutation-positive) or (3) are required to take potentially competing antidepressants.

5.3. PHASE II RESULTS

The study population was selected from 625 South African individuals enrolled in a chronic disease screening program, of whom 88 (14%) were previously diagnosed with breast cancer. The majority of these patients (53, 60%) reported a family history of cancer. Eighteen (20%) patients had advanced metastatic disease, of whom 14 (78%) had ER-positive carcinoma. A total of 264 individuals (194 females, 70 males; mean age 48.0 years) reported to be cancer-free at entry into the study were identified as obese. In addition, 118 obese patients (45%) had hypertension, 68 (26%) had hypoalphalipoproteinemia, 53 (20%) had hypertriglyceridemia and 49 (19%) had insulin resistance. In total, 94 obese patients had no other concurrent metabolic syndrome features, while 99 had one additional component. In total, 71 obese patients (27%) had concurrent metabolic syndrome, based on the presence of two or more additional features except visceral adiposity. Of the 214 female control individuals reported to be cancer-free and without a family history of cancer, 175 were classified as obese (BMI>30 kg/m2).

For comparative analysis among breast cancer patients, all cases were subdivided according to ER status. Histopathology reports were available for 81 (92%) of the 88 breast cancer patients participating in the chronic disease screening program. In addition to the information gathered from this patient group referred to as the implementation data set, ER status was also available for 83 breast cancer patients from the validation data set previously described by van der Merwe et al. (2012). HER2 status was not performed/ available for these patients, all of whom were previously screened for mutations in the BRCA1 and BRCA2 genes (Agenbag 2005). Table 5.7 summarises the clinical characteristics of the combined patient study group, including

104 (63%) Caucasian and 60 (37%) Coloured breast cancer patients (n=164, p-value=0.160) stratified according to ER status.

Table 5.7. Comparison of clinical and lifestyle characteristics between ER-positive and ER-negative breast cancer patients, adjusted for ethnicity and age.

Parameter		ER status		P-value, adjusted for:	
		ER-	ER+	Ethnicity	Ethnicity and age
		49	115		
Median (range)	**N**				
Age	164	49 (30-77)	54 (31-83)	**0.013**	
Age at diagnosis/onset	164	48 (30-77)	53 (31-83)	**0.022**	0.794
Body mass index	146	25 (17-47)	26 (17-41)	0.014	**0.035**
*Folate Score	81	8 (2-15)	8 (3-31)		0.232
*Fruit, veg and fibre score	81	10 (3-20)	14 (4-25)	All Caucasians	0.082
*Fat Score	81	11 (2-29)	12 (1-34)		0.865
Count (percentage)					
Ethnicity, Coloured	164	22 (45)	38 (33)		
HRT, yes	163	9 (19)	15 (13)	0.170	0.089
OCPs, yes	163	20 (42)	32 (28)	0.247	0.687
Current smoking, yes	160	17 (35)	34 (30)	0.898	0.831
Alcohol intake, yes	156	25 (52)	62 (57)	0.838	0.545
Physical activity, yes	162	33 (67)	74 (65)	0.427	0.392
Family history of cancer	163	30 (61)	59 (52)	0.164	0.371
BRCA mutation-positive	106	7 (19)	6 (9)	0.095	0.143

HRT, hormone replacement therapy; OCPs, oral contraceptive pills. *Only available for a subset of Caucasian patients i.e. participants of chronic disease screen (n=81), no ethnicity adjustment possible.

Within the two ethnic groups, there was no significant difference between the number of patients aged 50 years or older, used as a proxy for menopausal status. There were also no significant

differences between age (p=0.595) and age at diagnosis/onset of breast cancer (p=0.265) between Caucasian and Coloured breast cancer patients. However, modifiable lifestyle factors differed significantly between these two South African population groups as shown in Figure 5.10. BMI (p=0.0001) and current smoking (p<0.0001) were highest in the Coloured population, while level of physical activity (p=0.0001) and weekly alcohol consumption (p=0.0002) were highest in the Caucasian patients. Use of hormone replacement therapy (HRT) was significantly higher in Caucasian breast cancer patients compared to Coloured patients (p=0.0038), while use of oral contraceptive pills (OCP) was the highest in the Coloured breast cancer patients (p<0.0001).

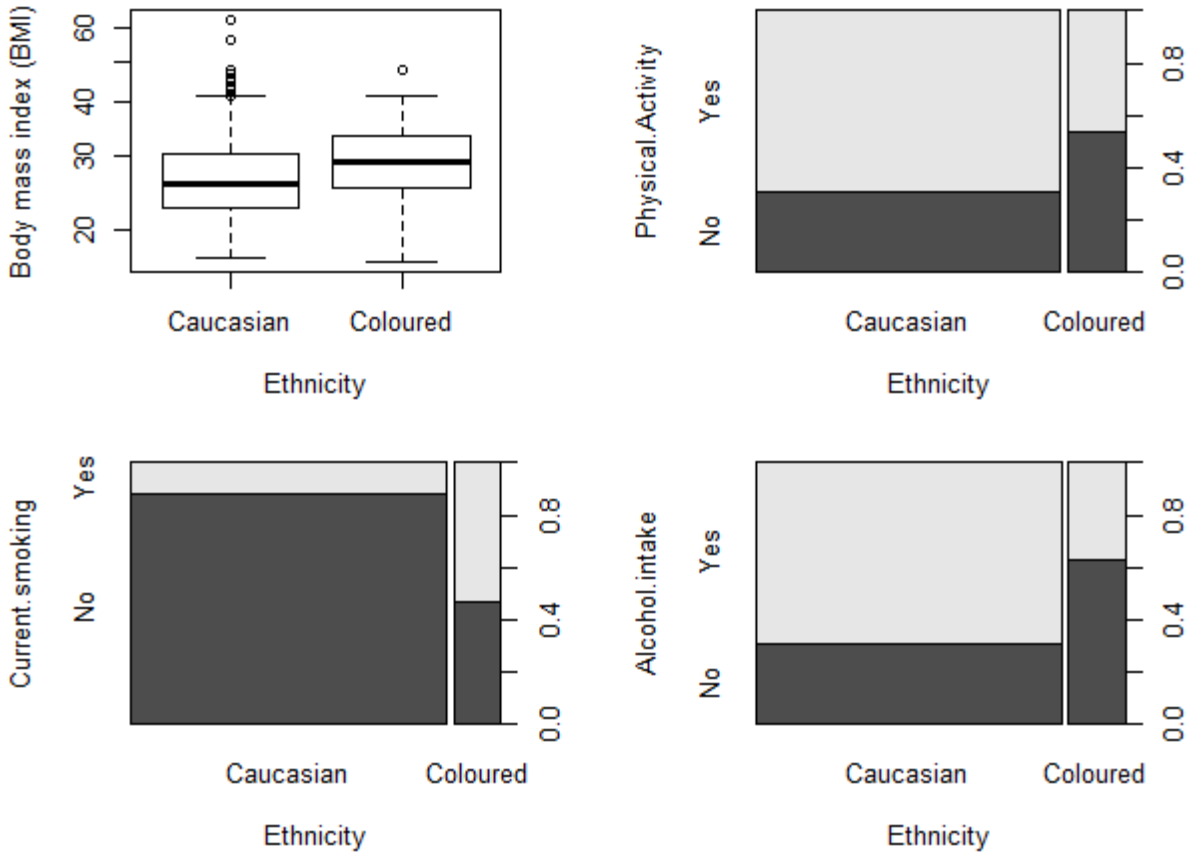


Figure 5.10. Comparison of modifiable lifestyle factors between breast cancer patients from the Caucasian and Coloured populations of South Africa.

All 87 patients included in the validation data set were previously screened for high-penetrance mutations in the entire BRCA1/2 genes (van der Merwe et al. 2012), while BRCA mutation analysis was performed in only 23 patients in the implementation data set. Selection criteria for BRCA mutation screening following genetic counselling in the latter group included early age at diagnosis (<40 years) and / or a strong family history of breast/ovarian cancer. Detection of

high-penetrance mutations in BRCA1 and BRCA2 genes were more commonly reported in patients with ER-negative (19%) compared to ER-positive (9%) tumours, as may be reflected by the higher proportion of patients in the former group who reported a family history of cancer. This finding also concurs with a significantly younger age at entry into the study (p=0.013), as well as a significantly earlier age at diagnosis/onset of breast cancer in patients with ER-negative versus ER-positive cancer, after adjustment for ethnicity (p=0.022). As indicated in Table 5.7, BMI was furthermore found to be significantly higher in patients with ER-positive compared to ER-negative breast cancer, after adjustment of age and ethnicity (p=0.035). Each extra unit of BMI increased the odds of being ER-positive with 8% (95% CI: 1-16).

Due to the high prevalence of lifestyle risk factors identified in the breast cancer patients studied, genotype-phenotype association studies were performed using a CVD multi-gene assay including 8 SNPs considered clinically useful across diagnostic boundaries (Table 4.7). These SNPs in the APOE, MTHFR, FII, FVL, and HFE genes are the most common hereditary causes for hypercholesterolaemia, hyperhomocysteinaemia, thrombophilia and iron overload, respectively, in the general population worldwide. The allelic distribution of all 8 SNPs were in Hardy-Weinberg equilibrium for the total and both subgroups of breast cancer patients studied (p>0.05). No significant differences in genotype distribution and allele frequencies were observed for the 8 SNPs included in the CVD multi-gene assay between patients with ER-positive and ER-negative breast cancer (Table 5.8).

Table 5.8. Comparison of genotype distribution and allele frequencies for 8 SNPs included in the CVD multi-gene assay between ER-positive and ER-negative breast cancer patients (112 Caucasian and 62 Coloured), after adjustment for ethnicity.

	Genotype		Additive allele	
Gene variant	**Unadjusted**	**Ethnicity-adjusted**	**Unadjusted**	**Ethnicity-adjusted**
ApoE2 C>T	0.195	0.234	0.195	0.234
ApoE4 T>C	0.143	0.198	0.150	0.242
MTHFR 677 C>T	0.644	0.747	0.570	0.904
MTHFR 1298 A>C	0.529	0.567	0.892	0.987
Factor II G>A	0.226	0.241	0.085	0.092
FVLL G>A	0.575	0.635	0.575	0.635
HFE 845 G>A	0.271	0.168	0.271	0.168
HFE H63D C>G	0.227	0.333	0.121	0.200

Minor allele frequencies for the 8 SNPs studied were similar to those previously reported for the general South African population, stratified by ethnicity (Table 5.9).

Table 5.9. Comparisons of minor allele frequencies (MAFs) of variants included in the CVD multi-gene assay applied in 112 Caucasian and 62 Coloured breast cancer patients in relation and MAFs of the respective SNPs according to the literature.

Gene variant	MAF Caucasian patients	MAF Caucasians (literature)	MAF Coloured patients	MAF Coloureds (literature)	GMX database and/or South African published data
APOE e-2	0.07	0.08	0.1	0.11	Kotze et al. 1993 and GMX database
APOE e-4	0.16	0.17	0.28	0.26	Kotze et al. 1993 and GMX database
FVL Leiden	0.04	0.05	0.02	0.02	GMX database
FII	0.02	0.02	0.01	0.02	GMX database
MTHFR 677 C>T	0.36	0.36	0.12	0.18	Scholtz et al. 2002
MTHFR 1298 A>C	0.25	0.37	0.32	0.3	Scholtz et al. 2002
HFE C282Y	0.04	0.09	0.03	0.02	de Villiers et al. 1999
HFE H63D	0.15	0.12	0.125	0.07	de Villiers et al. 1999

GMX database = Gknowmix database accessible to registered users at www.gknowmix.org

Extended genotype-phenotype association studies demonstrated (Figure 5.11) that presence of the E4 allele of the ApoE polymorphism (rs429358) reduces the mean age of breast cancer diagnosis/onset by an average of 10 years in South African breast cancer patients, regardless of ethnicity, alcohol consumption, current smoking and a family history of cancer (P=0.003).

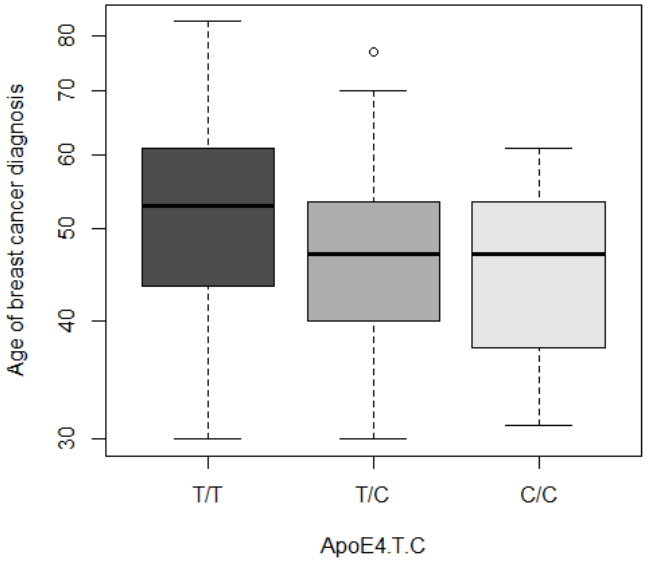


Figure 5.11. Box plot depicting the additive allelic reducing effect of APOE allele E4 on age of onset/diagnosis of breast cancer patients.

The functional MTHFR 1298 (A>C) polymorphism (rs1801131) was furthermore found to be significantly associated with BMI, after adjustment for ethnicity and a family history of cancer (p=0.01), as shown in Figure 5.12.

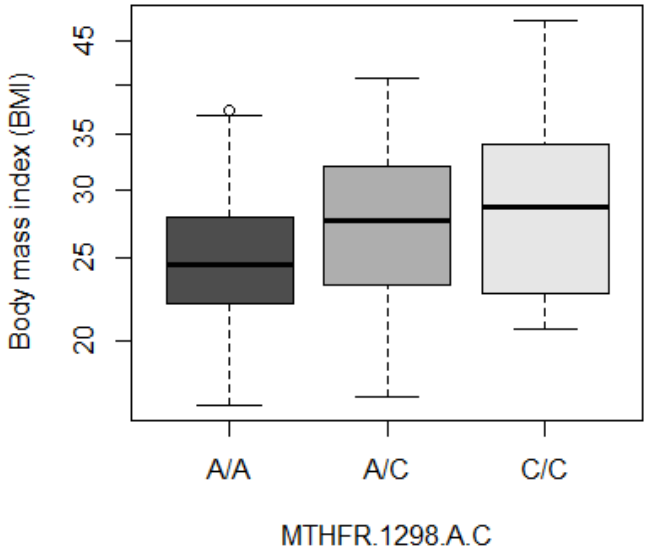


Figure 5.12. Box plot depicting the additive allelic increasing effect of MTHFR 1298 A>C on body mass index (BMI).

Association analysis was performed between body mass index (BMI) and the dietary intake in 214 female Caucasian control individuals and 81 breast cancer patients from the implementation data set who completed the nutrition questionnaire (Figure 5.13). A difference in association was detected between the patient and control groups (p=0.049). The patient group showed a significant increasing trend of BMI with higher saturated/trans fat score (p=0.032), whereas the same association was not found in the control group. When the association between BMI and the fat score was assessed in the implementation data set, separated by ER status, no significant difference was observed between the two breast cancer patient groups. The positive correlation between BMI and the fat score was stronger in ER-positive compared to ER-negative breast cancer patients, but the sample size may have been too small to reach statistical significance.

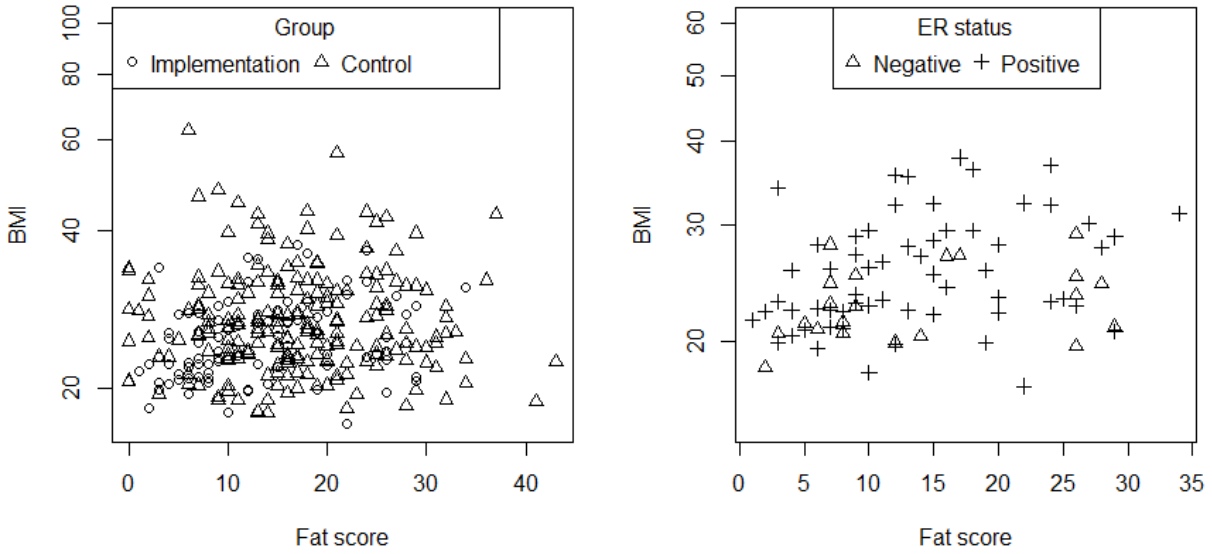


Figure 5.13. Association between BMI and saturated/trans fat score. Left: Comparative analysis in 214 female Caucasian control individuals and 81 breast cancer patients from the implementation data set who completed the nutrition questionnaire, showing a difference in association between groups (p-value=0.049). Right: Association between BMI and fat score in implementation data set, separated by ER status.

Seventeen patients included in the implementation data reported treatment-related side effects, including chemotherapy-induced cognitive dysfunction and/or drug resistance/side effects related to the concomitant use of tamoxifen/aromatase inhibitors. In addition, four patients reported side effects with the use of aromatase inhibitors and other drugs. Genotype-phenotype correlation of these 17 cases showed that in eight patients, the clinical profiles matched their

genetic test results. Four (4.5%) breast cancer patients reported symptoms of cognitive impairment, two of whom tested positive for both the risk-associated APOE e-4 allele and the MTHFR 677 C>T mutation previously linked to this condition. A total of 27 (30%) breast cancer patients reported a comorbid diagnosis of depression, 12 (44%) of whom were MTHFR 677 T-allele carriers. Of the nine patients who tested positive for mutations in the FVL Leiden/prothrombin gene, four MTHFR 677C>T carriers (out of 7 patients) also experienced deep vein thrombosis (DVT) and/or recurrent pregnancy loss (Table 5.10).

Table 5.10. Clinical characteristics and comorbidities in breast cancer patients enrolled in the chronic disease screen, subjected to the CVD multi-gene assay.

Comorbidity	Breast cancer patients (n=88)
BRCA mutation analysis (mutation positive)	23 (10)
Depression (%)	27 (31%)
Depression (antidepressant treatment) (%)	10 (11%)
Medication side effects (%)	17 (19%)
Tamoxifen use	7
Antidepressant use	5
Tamoxifen + Antidepressant	1
Other including AIs	4
Deep vein thrombosis (%)	2 (2%)
Recurrent pregnancy loss (%)	5 (5%)
Cognitive Impairment (%)	4 (4%)

5.4. PHASE II DISCUSSION

Comparison between patients with ER-positive and ER-negative breast cancer showed a statistically significantly higher BMI in hormone-responsive patients after adjustment for age and ethnicity (p=0.035). Differences in several modifiable lifestyle factors were observed between Caucasian and Coloured breast cancer patients, including BMI found to be significantly higher in the Coloured patients. The MTHFR 1298 A>C SNP was significantly associated with increased BMI (p=0.01) in the present study, after adjustment for ethnicity and a family history of cancer. This finding may be related to the association reported by Naushad et al. (2012) linking variation

in the MTHFR gene with the luminal B tumour subtype. Two of the 13 South African patients who were HER2 positive were found to be homozygous for MTHFR 677 C>T; both ER-positive and therefore classified as luminal B as opposed to the HER2-enriched subtype characterised by ER-negative status.

The E4 allele of the apolipoprotein E (APOE) polymorphism (rs429358) reduced the mean age of cancer diagnosis/onset by an average of 10 years, regardless of the aforementioned confounders (P=0.003). The APOE e-4 allele included in the CVD multi-gene assay previously described by Kotze and Thiart (2003) has previously been associated with breast cancer risk and clinical outcome across different population groups (Zunarelli et al. 2000; Moore et al. 2004; Chang et al. 2005).The APOE gene provides a genetic link between CVD and dementia, with the cholesterol-raising APOE e-4 allele considered an important risk factor for metabolic syndrome, vascular disease and cancer. Women with one or two copies of the e-4 allele of APOE and those with high concentrations of triglycerides had a four-fold risk of developing breast cancer, compared to women with low triglyceride levels. This increased risk was suggested to be attributed to factors including the effect of reduced triglyceride clearance from the plasma which results in sustained elevated concentrations and consequently a decrease in the levels of sex hormone-binding globulins (Moysich et al. 2000). This is in agreement with the findings of Chang et al. (2005), Surekha et al. (2008) and a meta-analysis performed by Saadat (2012) in 2008 breast cancer patients and 1 827 healthy controls. Saadat (2012) concluded that the APOE e-4 allele may be a low-penetrant risk factor for development of breast cancer. Other studies however reported no association between the presence of either the APOE e-4 allele and the rate of tumour cell proliferation or clinical outcome in breast cancer patients (Zunarelli et al. 2000). The deleterious effect of the APOE e-4 allele is not only exerted through increased cholesterol levels, but also independently through increased risk of LDL-oxidation and inflammation. Genin et al. (2011) demonstrated the association of APOE e-4 with early-onset Alzheimer's disease (AD) and calculated AD life-time risk in approximately 10 000 cases and controls of Caucasian ancestry. This risk increased with the presence of the APOE e-4 from 11% to 51% in males and 14% to 60% in female carriers at the same age. They contended that similar to BRCA gene mutations, the aforementioned risks when compared to that conferred by low-risk common alleles, are consistent with the semi-dominant inheritance of moderately penetrant genes.

Lengacher et al. (2015) studied chemotherapy-related changes in cognitive function in breast cancer survivors participating in the Mindfulness-Based Stress Reduction (MBSR) for Breast Cancer program. The results showed that variation in four genes including APOE and MTHFR were associated with cognitive impairment. In APOE e-4 allele carriers, co-inheritance of the

MTHFR 677C>T variant has been shown to exacerbate risk for cognitive dysfunction as a known chemotherapy-induced side-effect. In a recent South African study conducted by Luckhoff et al. (2015) the authors demonstrated that the expression of a hypercholesterolaemic phenotype in e-4 allele carriers, as well as its mitigation by regular physical activity, were dependent on the interaction between APOE genotype and AD family history. In an extended data set which included all participants enrolled in this study, APOE genotype was also shown to modify the association between alcohol intake and metabolic risk traits including BMI and total cholesterol levels, with a deleterious increase in both being restricted to APOE e-4 allele carriers (unpublished data).These findings served to validate the clinical relevance of assessing non-genetic data including lifestyle factors as part of a multidisciplinary approach to chronic disease screening incorporating APOE genotyping to inform clinical and therapeutic decision making in the context of vascular risk management. APOE genotyping could therefore prove useful in prioritizing the need for tailored lifestyle-based interventions in e-4 allele carriers with the goal of decreasing cumulative cardio-metabolic risk to prevent the onset and progression of CVD/AD with advancing age. This emphasizes that the clinical application for APOE genotyping performed as part of a multidisciplinary framework is not applied to provide diagnostic confirmation of a fatalistic diagnosis, but as a means of developing and providing lifestyle-based treatment plans aimed at overcoming an “apparently insurmountable” genetic disease predisposition (Kotze et al. 2015).

In this study, four breast cancer patients reported symptoms of cognitive impairment at the time of enrolment into the chronic disease screening program, two of whom tested positive for both the risk-associated APOE e-4 allele and MTHFR 677 C>T variant (heterozygous and homozygous, respectively) and another heterozygous for MTHFR 677 C>T. One of the patients who were homozygous for MTHFR 677 C>T and heterozygous for APOE e-4 was ER-positive while the patient heterozygous for MTHFR 677 C>T was ER-negative (triple-negative). Co-inheritance of these two functional polymorphisms or single effects in combination with relevant lifestyle risk factors detected with use of the CVD assay may therefore explain the cognitive changes observed in a subset of breast cancer patients following anti-cancer treatment. This was in accordance with a meta-analysis conducted by Jim et al. (2012) in 807 patients previously treated with standard-dose chemotherapy showing that deficits in cognitive functioning were observed in patients relative to controls or pre-chemotherapy baseline results with regard to verbal ability and visuospatial ability. In addition, patients treated with chemotherapy performed worse than controls in relation to verbal ability and worse than patients treated without chemotherapy with regard to visuospatial ability. Possible confounders including age, education, endocrine treatment and duration since treatment did not moderate cognitive deficits in verbal ability or visuospatial ability. A study by Stewart et al. (2006)

supported these results. Continued monitoring of patients through our combined PSGT service and research approach enables the identification and investigation of treatment-related comorbidities. This approach proved useful for breast cancer patients subjected to the CVD multi-gene assay covering all major disease pathways implicated in cancer and other non-communicable diseases.

Results from numerous, large-scale prospective cohort studies confirmed the increased risk of CVD mortality with increasing BMI (McGee 2004; Flegal et al. 2005; Wilson et al. 2002). Several studies have also reported an association between obesity as well as weight gain and breast cancer risk (Wolket al. 2001; Huang et al. 1997). There is also evidence that weight loss as well as reduced fat intake may lead to a decreased risk for breast cancer (Harvie et al. 2005; Prentice et al. 2006). In addition to being an established risk factor for CVD, a direct association between metabolic syndrome and postmenopausal breast cancer risk has been clearly demonstrated (Agnoli et al. 2015, Rosato et al. 2011, Davis et al. 2012). Metabolic syndrome is also a common comorbidity in patients with depression, which according to various studies affects about 30% of breast cancer patients (Vahdaninia et al. 2010; Saniah et al. 2010). This finding is in agreement with the prevalence of depression in our study group representing an expansion of the “validation data set” previously studied by van der Merwe et al. (2012a). In contrast to a bi-directional pathogenic relationship, it has been suggested that a consistent overlap between common metabolic abnormalities and neuropsychiatric features shared across the affective spectrum could support the existence of a distinct metabolic-mood syndrome disease subtype. The association between cardio-metabolic risk and depression can also be ascribed to side effects related to common antidepressants, including weight gain which was reported by one patient included this study. These findings emphasize the importance of early detection of a genetic predisposition to allow for classification into different treatment groups and enable the implementation of measures to prevent cumulative risk that may underlie the development of breast cancer and associated comorbidities.

The MTHFR 677 C>T mutation was previously shown to be associated with increased BMI in South African patients with major depressive disorder (Delport et al. 2014). The clinical validation of MTHFR 677 C>T as a contributing factor to cumulative metabolic risk was demonstrated in the South African population by the interrelationship detected between dietary folate intake, MTHFR mutation status, BMI and homocysteine levels (Delport et al. (2014). These findings are supported by an earlier study of Tiemeier et al. (2002) who also showed that the effect of folate deficiency was mediated by obesity, which was identified as a major comorbidity in patients with depression. Our data supported by international studies therefore provides a scientific basis for personalized dietary recommendations based on an individual's

genetic makeup and nutritional status that may be reflected by relevant biochemical abnormalities. Since depression is the most common psychiatric problem in cancer patients, treatment focused on both conditions using a comprehensive risk reduction approach guided from the genetic background has become an important consideration in the local population (van der Merwe et al. 2012a). Up to 30% of patients treated with tamoxifen to reduce recurrence risk also use anti-depressants (for depression or hot flushes) (Nelson et al. 2006; Kim et al. 2010) which may interfere with anti-cancer treatment.

In our study, one breast cancer patient also identified as being homozygous for the MTHFR 677C>T variant (BRCA mutation negative) presented with side effects attributed to tamoxifen use which could not be accounted for by the presence of inactivating CYP2D6 alleles. This patient was subjected to gene expression profiling using the MammaPrint test which revealed the presence of bilateral low-risk breast cancer. The daughter of this patient also developed breast cancer approximately two years after her mother. Based on the clinical presentation and assessment of her genetic risk profile, this patient was selected for extended genetic testing in the form of next-generation/whole exome sequencing.

Co-inheritance of MTHFR 677C>T with high-penetrance causative mutations in the BRCA1/2 genes may also increase the risk of early-onset breast and ovarian cancer (Chen et al. 2005; Jakubowska et al. 2007; Pepe et al. 2007). In the South African setting, elucidation of the role of MTHFR as a modifier gene in familial breast cancer is of particular relevance, given the detection of founder BRCA1/2 mutations at increased frequencies in the Afrikaner, Coloured and Xhosa population groups (Reeves et al. 2004; van der Merwe et al. 2012b). Similar to the low-penetrance MTHFR 677C>T mutation, distribution of BRCA1/2 mutation status was the same in ER-positive and negative patients included in this study. This finding supports the relevance of susceptibility genetic testing across the spectrum of breast cancer subtypes, ranging from the luminal-type associated with variation in the MTHFR gene (Naushad et al. 2012) to the more severe HER2 enriched and basal-like triple negative breast cancers strongly associated with BRCA1 mutations. The distribution of HER2 overexpression was similar in patients with ER-negative (20%) and ER-positive (15.5%) cancers. Of the 10 (5 ER-positive and 5 ER-negative) patients who tested positive for mutations in the BRCA1 and BRCA2 genes, at least one patient with ER-positive breast cancer was previously found to be homozygous for the inactivating CYP2D6 allele 4 allele as highlighted in the study by van der Merwe et al. (2012a) due to a change in treatment based on the genetic findings. Extension of CYP2D6 genotyping to using the CVD multi-gene assay in the same patient revealed homozygosity for the MTHFR 677 C>T mutation as well as homozygosity for the APOE e-4 allele, both considered established risk factors for Alzheimer's disease and susceptibility for cancer development. Another example

of how early intervention was enabled by participation in the chronic disease risk screen was demonstrated by the case of a healthy but obese 51 year old female encouraged to have genetic counselling and BRCA testing based mainly on her family history of cancer documented as part of the medical and lifestyle assessment. She requested re-evaluation of a recent mammogram performed that led to further testing and finally a diagnosis of early-stage triple negative (ER/PR/HER2-negative) breast cancer. Subsequent full gene BRCA screening as part of the MyRisk test performed by Myriad Genetics, resulted in the identification of the BRCA1 c.2457del mutation. This information has now been added to our genomics database after exchange from “control” to “case” in order to ensure the accuracy of the information made available for future research. In relation to cancer diagnosis two other patients included in this study reported a change in disease status during the 5-year study period. In our study, one breast cancer patient also identified as being homozygous for the MTHFR 677C>T variant (BRCA mutation negative) presented with side effects attributed to TAM use which could not be accounted for by the presence of inactivating CYP2D6 alleles. This patient was subjected to gene expression profiling using the MammaPrint test which revealed the presence of bilateral low-risk breast cancer. The daughter of this patient also developed breast cancer approximately two years after her mother. Based on the clinical presentation and assessment of her genetic risk profile, this patient was selected for extended genetic testing in the form of next-generation/whole exome sequencing.

The selection of this patient for next-generation sequencing was based on the need to establish whether MTHFR 677C>T homozygosity would be sufficient to confer significant genetic risk for breast cancer. Results from such investigation could help to test the hypothesis that low-penetrance functional polymorphisms in genes such as MTHFR are associated with genetic predisposition towards breast cancer in the presence of a high-risk environment including an increased BMI. It is important to determine whether stop-codon mutations in genes such as RAD50 identified for these patients should be considered causative in the presence of non-deterministic genetic risk modifiers such as MTHFR 677C>T. This notion is in accord with findings from a recent study in which the authors showed that the majority of familial aggregation for breast cancer can be attributed to low- to- moderate-penetrance mutations in non-BRCA1/2 mutated cases.

The finding that tamoxifen use is associated with increased risk of thromboembolic events (TEs) in women with breast cancer is of particular relevance to patients with a genetic predisposition for DVT. Although our sample size is too small to obtain conclusive results, it is noteworthy that none of the patients with ER-negative breast cancer has the FVL Leiden or prothrombin mutations detected in 10% and 5% of patients with ER-positive tumours, respectively. In a study

by Garber et al. (2010) studying 124 women who had experienced TEs while taking adjuvant tamoxifen (±5 years) for breast cancer, a significant association with the FVL 1691G>A (Leiden) mutation was identified. Other statistically significant factors associated with TE risk were personal history of TE and smoking also captured as part of the medical and lifestyle assessment performed as part of the chronic disease screening program. Since women who experienced TEs while taking adjuvant tamoxifen for early-stage breast cancer were five times more likely to carry the FVL Leiden mutation than those who did not have a TE, it was suggested that postmenopausal women should be evaluated for the FVLL mutation before prescription of adjuvant TAM if a positive test would alter therapeutic decision making.

The CVD multi-gene assay used in this study as the genetic component of the chronic disease screening program includes both the FVL and FII / prothrombin mutations (Kotze and Thiart 2003). These low-penetrance mutations are the two major coagulation factors essential for proper blood clot formation and are the most common inherited risk factors for venous thrombosis. Risk for thrombosis in carriers of these mutations is further increased with the use of estrogen-containing oral contraceptives (Swen et al. 2011). This pharmacogenetic interaction underpins the importance of FVL as an established risk factor for CVD, also in breast cancer comorbid risk assessment. Due to differences reported in the prevalence of venous thromboembolism in major population groups, Pepe et al. (1997) determined the distribution of the FVL mutation in European and various non-European groups. They confirmed that in non-Europeans, the prevalence of FVL mutation is at least 7 times lower than in Europeans, which may explain the low prevalence of thrombotic event in these groups. MTHFR 677 C>T has also been documented as a risk factor for venous and arterial thrombosis and the risk may be further increased in the presence of the FII and FVL mutations (Elhassan and Abdalla 2015; Andreassi et al. 2006; Den et al. 2005). Keijzer et al. (2002) reported a relative risk for thrombophilia of between 1.4 and 1.6 in the presence of the MTHFR 677 T-allele and a combined risk of 18.7 if detected in the presence of the FVL Leiden mutation. The development of thrombosis has been associated with various clinical conditions including arterial thrombosis (myocardial infarction, ischemic cerebrovascular disease, and peripheral vascular disease) and venous thrombotic disorders [deep vein thrombosis (DVT) and pulmonary embolus]. Furthermore, both the APOE e-2 and e-4 alleles included in the CVD multi-gene assay have been associated with DVT (Nagato et al. 2012), supporting the relevance of this gene in both CVD and breast cancer risk management. Finally, the APOE e-2, APOE e-4, FVL, FII and MTHFR 677 C>T polymorphisms have been associated with recurrent pregnancy loss (Reznikoff-Etiévan et al. 2001; Mierla et al. 2012; Li et al. 2014). Of the nine patients who tested positive for the FII and FVL gene variants in our study, four also carried the MTHFR 677 C>T variant and reported a history of prior DVT and/or recurrent pregnancy loss. This result provides a good example of concordance between

the genetic test result and clinical expression, which supports the clinical relevance of using the CVD multi-gene assay in breast cancer patients as part of a comprehensive chronic disease screening program.

In addition to its anti-neoplastic effects, several clinical studies have demonstrated that tamoxifen exerts favourable effects on serum lipid profiles and metabolic risk, being associated with lower levels of total and low-density lipoprotein (LDL)-cholesterol (de Medina et al. 2004). Cholesterol which is a precursor for estrogen production, is packaged into LDL which when inhibited with the use of statins, was shown to reduce the incidence and increase patient survival (Beck et al. 2003; Cauley et al 2006). In light of the CVD risk faced by breast cancer survivors, lowering of cholesterol levels has been suggested as therapy for targeting both diseases. The alternative to lowering cholesterol levels without use of medication is dietary intervention. Various diets have been investigated for their potential effect on cancer risk, with conflicting results. Toledo et al. (2015) suggest a beneficial effect of a Mediterranean diet (MeDiet) supplemented with extra-virgin olive oil (EVOO) in the primary prevention of breast cancer. In this study, women allocated to the MeDiet supplemented with EVOO showed a 62% relatively lower risk of malignant breast cancer than those allocated to the control diet (95% CI, 0.16-0.87). While HRT was associated with favourable changes in lipid and metabolic parameters in carriers homozygous for the C-allele of MTHFR 677, this effect was not evident in carriers of the risk-associated T- allele. The MTHFR C677T polymorphism was suggested to modify the effect of HRT on lipid and metabolic parameters in postmenopausal women (Lambrinoudaki et al. 2013). The homozygous mutant TT genotype was associated with increased levels of total and LDL cholesterol, while the heterozygous CT and wild type CC genotypes were associated with a reduction of total cholesterol and LDL cholesterol after 1 year of HRT ($p = 0.032$ for total cholesterol and $p = 0.002$ for LDL cholesterol). Devlin et al. (2012) suggested that the high prevalence of MetS as a side-effect of treatment with second generation antipsychotics in children, is attributed to the role of the MTHFR C677T variant in the metabolic alterations that link these two constructs.

Mutations in the hemochromatosis (HFE) gene also included in the CVD multi-gene assay account for up to 90% of type I hereditary hemochromatosis (HH) in patients of Northern European descent (Adams and Barton 2007). Variation in the HFE gene is associated with increased risk for iron overload as well as other related comorbidities including cardiomyopathy and chronic liver disease (Gurrin et al. 2008). In addition, iron as an essential trace element can be carcinogenic through a variety of mechanisms including host immune response suppression (Porto and De Sousa 2007), catalysing the formation of mutagenic hydroxyl radicals (Ames et al. 1983) and by stimulating growth of tumour cells (Stevens et al. 1990). Iron metabolism was

shown to be frequently altered in breast cancer, with considerably higher iron in postmenstrual women. Moreover, a synergistic effect on cellular proliferation between estrogen and iron has been demonstrated in BRCA1-associated hormone responsive breast cancer (Mittal et al. 2014). Zghair et al. (2014) contended that establishing a link between estrogen and iron metabolism has a potentially significant prognostic value in predicting clinical outcome in BRCA1-associated ER-positive breast cancer patients. Kallianpur et al. (2004) concluded that impaired iron homeostasis in HFE C282Y mutation carriers predisposes towards the development of breast cancer in Caucasian females. In a prospective cohort study, Osborne et al. (2010) assessed the relationships between the low-penetrance HFE C282Y mutation and risk for the development of several neoplasms including breast, colorectal, and prostate cancers in nearly 30 000 participants. Males homozygous for the C282Y mutation had an increased risk of colorectal cancer compared to non-carriers, while female mutation carriers had an increased risk for breast cancer. Overall, HFE C282Y homozygotes had a two-fold increased risk of breast and colorectal cancer compared with those who tested negative for the C282Y. None of the breast cancer patients included in the present study were homozygous for the C282Y mutation.

The small sample size and use of age as a proxy for menopausal status limited our ability to replicate the findings of Lambrinoudaki et al. (2013), who identified MTHFR 677C>T as a modifier of the effect of HRT on metabolic risk in postmenopausal females. Since accurate designation of menopausal status is complicated by various factors including surgery and the use of hormone therapy, we based our decision to use an age of 50 years and older as in indication of postmenopausal status on the findings of Phipps et al. (2010). These authors reported that different methods used to define menopausal status translated to only slight differences in status-specific breast cancer incidence or detection rates. Inclusion of only Caucasian patients of European Ancestry is another important limitation of the study, which can be justified by the low frequencies of most of the SNPs included in the CVD multi-gene assay in non-Caucasians. Homozygosity for the MTHFR 677 T-allele was reported to be approximately 10% in the Caucasian population, while virtually absent in in the Black population of South Africa (Scholtz et al. 2002) and reported to be underrepresented in Sub-Saharan Africa (Bauduer and Lacombe 2005). The frequency of the MTHFR 677 T-allele was approximately 7% in South African patients of Indian descent correlating with an increased risk for coronary artery disease (Ramkaran et al. 2015). Significant population differences were also reported for SNPs analysed in the FII, FVL and HFE genes (Hira et al. 2003; de Villiers et al. 1999), while the risk implications of the APOE polymorphism found to be the same or higher in non-Caucasian populations may be different in Africans (Luckhoff et al. 2015). The strength of this study however lies in the description of the added value of facilitating the clinical application of the CVD multi-gene risk assay in patients with breast cancer to concurrently assess risk for

relevant comorbidities with a particular emphasis on metabolic syndrome and vascular risk. In addition, this study was conducted in an uncontrolled research setting, which limited the potential for pre-selection of a specific disease phenotype.

In this study we demonstrated the potential value of the CVD multi-gene assay positioned alongside the assessment of relevant non-genetic data to guide clinical and therapeutic decision making across diagnostic boundaries. A multi-disciplinary approach to chronic disease risk management could assist clinicians in the development and timely implementation of personalized lifestyle-based intervention strategies aimed at decreasing cumulative cardio-metabolic risk in patients with or at risk for breast cancer. In addition to its added value in the context of chronic disease risk screening, the clinical application of our PSGT approach could also assist clinicians in the identification of patients set to derive optimal benefit from extended genetic testing including BRCA screening and ultimately whole genome sequencing in genetically uncharacterised patients or those experiencing drug side effects or failure. A genomic solution is provided to facilitate the lowering of cumulative risk and at the same time identify the need for further analysis beyond the limitations of the CVD multi-gene assay.

5.5. PHASE III RESULTS

WES was performed in a total of three breast cancer patients and three controls selected for comparison, using the newly developed exome prescreen algorithm (EPA) as illustrated in Figure 5.14.

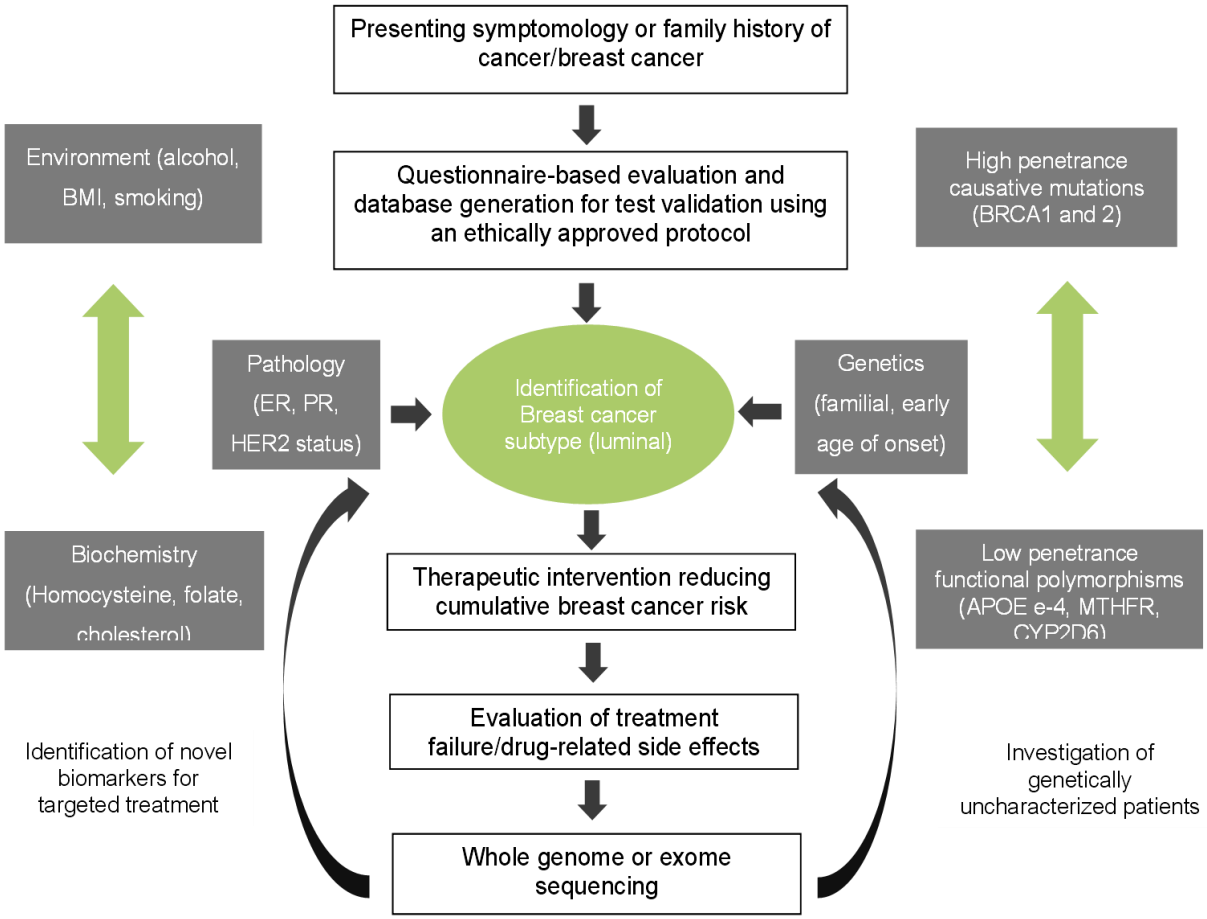


Figure 5.14. Outline of PSGT utilizing an open-innovation platform in order to facilitate targeted treatment as well as identify genetically uncharacterized patients eligible for whole exome sequencing. Adapted with permission from Kotze et al. (2015).

5.5.1. Discordance between Exome variant calling using different reference genomes

Filtering of variants on the minimum phred quality score and coverage resulted in a discordance rate of >20% in the index case (Figure 5.15) and unrelated breast cancer patient and controls (Table 5.15). Alignment of the exome sequence reads to hg19 resulted in many more variants detected compared to the MARS.

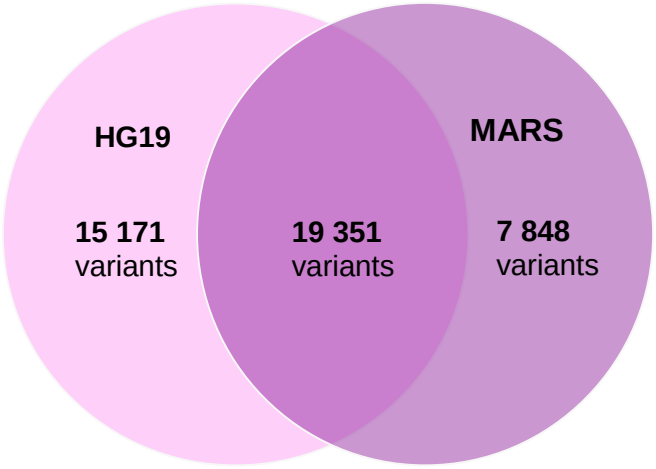


Figure 5.15. Venn diagram depicting the discordance between variants (filtered on a Q score of 50 and coverage of 100x) detected with read alignment to hg19 (53 743 in total) compared to CEU-MARS (79 153 in total), using low-stringency variant call settings.

Table 5.11. Comparison of variants obtained using hg19 versus the MARS in the unrelated breast cancer patients and three controls.

Exome	hg19 variants	MARS variants	Overlapping variants	Discordance
Unrelated patient	20 153	15 383	11 795	24%
Control 1 (1908)	21 676	16 569	12 437	24%
Control 2 (2550)	22 571	17 855	13 057	21%
Control 3 (3711)	20 632	16 142	12 264	22%

5.5.2. BRCA1 and BRCA2 mutation screening

Evaluation of WES results obtained for the BRCA1 and BRCA2 genes in the index patient with familial breast cancer and the unrelated patient with sporadic TNBC confirmed the absence of three high-penetrance founder mutations responsible for more than 80% of BRCA mutations in the Afrikaner population of South Africa: BRCA1 1493delC, BRCA1 2760G>T (p.Glu881X) and BRCA2 8162delG (Reeves et al. 2004; Van der Merwe and van Rensburg 2009). In addition, three Ashkenazi Jewish founder mutations (BRCA1 185delAG, BRCA1 5382insC and BRCA2 6174delT) as well as a recently described mutation (BRCA2 5999del4) found to be more common in patients of Xhosa and Mixed Ancestry in the Western Cape region of South Africa were not detected in these two patients.

Fifteen BRCA1 and nine BRCA2 SNPs were previously identified in the index patient at a commercial laboratory that performed targeted gene sequencing for familial breast cancer combined with multiplex ligation dependent probe amplification (MLPA) to detect deletions and insertions in BRCA1 and 2 genes (NewGene, UK). These included a BRCA2 missense variant (rs144848) with a minor allele frequency (MAF) of 0.25 identified in the index patient, but not in her mother nor in the unrelated breast cancer patient. This SNP was predicted to have no deleterious or damaging effect on the resulting protein according to PROVEAN and SIFT. Another BRCA1 SNP (rs799917) (MAF=0.46) identified in this patient was predicted to have no deleterious or damaging effect on the protein function, according to PROVEAN and SIFT. One of the BRCA1 polymorphisms (rs16941) (MAF=0.34) was predicted to have a potentially deleterious/damaging effect on the resulting protein according to PROVEAN and SIFT. This prediction however did not correspond with the result of Clinvar. An additional SNP in the BRCA1 gene (rs1799966) (MAF=0.36) was predicted to be benign using Clinvar and PROVEAN, while predicted to be potentially damaging by SIFT. These two single nucleotide polymorphisms were not identified in the mother of the index patient simultaneously subjected to WES. Seven BRCA1 and four BRCA2 SNPs were identified in the unrelated TNBC patient, including the aforementioned BRCA1 variants rs16941, rs799917 and rs1799966. The remaining BRCA1/2 variants identified by WES were predicted to be neutral by all three analytical prediction tools used. Table 5.12 lists the BRCA1 and BRCA2 variants detected by WES aligned to the MARS in the unrelated breast cancer patient and indicates which variants are shared with the family.

Table 5.12. BRCA variants identified in the triple negative breast cancer patient, using WES aligned to the MARS

Chr position	rs number	Ref/ alternate allele	Gene	Clinical significance (NCBI Clinvar)	Function prediction (PROVEAN)	Function prediction (SIFT)
17: 41223094***	rs1799966	T,C	BRCA1	Benign/ conflicting	Neutral	Damaging
17: 41234470	rs1060915	A,G	BRCA1	Benign	Neutral	Tolerated
17: 41244000***	rs16942	T,C	BRCA1	Benign	Neutral	Tolerated
17: 41244435***	rs16941	T,C	BRCA1	Benign	Deleterious	Damaging
17: 41244936**	rs799917	G,A	BRCA1	Benign	Neutral	Tolerated
17: 41245237**	rs16940	A,G	BRCA1	Benign	Neutral	Tolerated

17: 41245466	rs1799949	G,A	BRCA1	Benign	Neutral	Tolerated
13: 32911888	rs1801406	A,G	BRCA2	Benign	Neutral	Tolerated
13: 32912299	rs543304	T,C	BRCA2	Benign	Neutral	Tolerated
13: 32913184	.--	A,G	BRCA2	--	Neutral	Tolerated
13: 32915005	rs206076	G,C	BRCA2	Likely Benign	Neutral	Tolerated
13: 32906729*	rs144848	A,C	BRCA2	Benign	Neutral	Tolerated

* SNP present only in index patient;
** SNP present in both the index patient and unrelated breast cancer patient
*** SNP present in all three breast cancer patients

Using the Integrative Genome Viewer (IGV) software tool for visualization of the exome-sequenced data, we observed high coverage along the exomes of the BRCA1 and 2 genes (Figures 5.16 and 5.17).

Figure 5.16. Snapshot of coverage of the BRCA1 gene for the three exomes using the Integrative Genome Viewer (IGV).

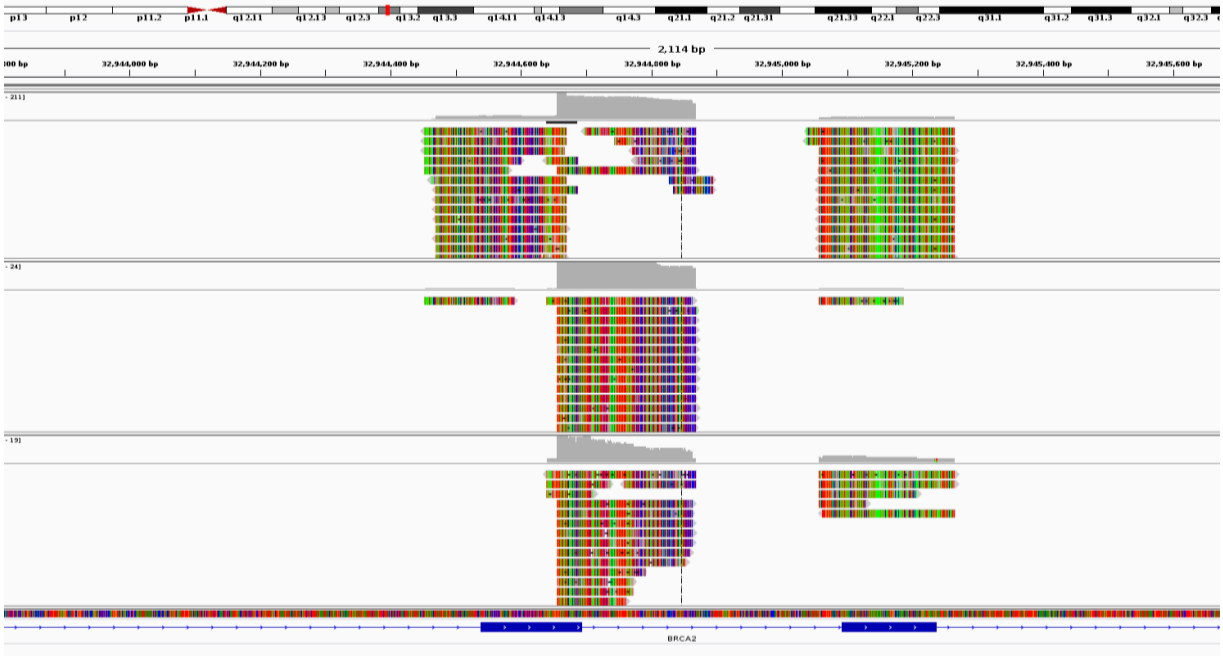


Figure 5.17. Snapshot of coverage of the BRCA2 gene for the three exomes using the Integrative Genome Viewer (IGV).

5.5.3. Comprehensive cancer panel screen using WES

Subsequent to searching for mutations in the BRCA1/2 genes using WES to determine the performance of the WES platform used in comparison with targeted BRCA1/2 sequencing previously performed in the index case, we shifted our focus to identification of rare variants of high-intermediate penetrance. This was followed by evaluation of common low-penetrance mutations involved in DNA mismatch repair as potential pharmacogenetic targets.

The analysis of variants included in the Ion AmpliSeq™ Comprehensive Cancer Panel (filtered on a coverage of 100x and phred score of 50) identified nineteen variants in the index patient, five of which were shared with her mother. Compared to the sequence quality of the index patient (minimum phred score and coverage of 548 and 291, respectively), her mother's minimum phred score and coverage (28 and 16, respectively) for shared variants were substantially lower. This low coverage and quality scores prompted use of the integrative genomics viewer (IGV) software tool for confirmation of the variants detected by WES. Subsequent to viewing the index patient and her mother's BAM files [binary version of a SAM (Sequence Alignment/Map) file] in IGV for investigation of the gene variants, we selected 7 apparently true variants for further analysis.

These 7 gene variants include two PDE4DIP variants (rs11295415 and rs11268079), MYH9 (chr22:36717864), MUC1 (chr1:155160668), two IGF2R variants (chr6:160471661 and chr6:160471662) and RAD50 (rs139372231) (Table 5.11). The RAD50 missense mutation (rs139372231) (MAF=<0.001) (http://exac.broadinstitute.org/variant/5-131924480-C-T) was documented as an intermediate-risk variant for breast cancer. IGV confirmed heterozygosity of RAD50 rs139372231 in the family but not in the unrelated breast cancer patient and 3 controls. Using IGV, a few variants including PDE4DIP (chr1: 144873963), EPHA (chr6: 94120780) and KMT2D (chr12: 49444818) were confirmed to be false positives. These as well as other gene variants appeared to be either artifacts, or not associated specifically with breast cancer based on a comprehensive literature search. They were therefore excluded from extended analysis and verification using Sanger sequencing/ Taqman genotyping.

Table 5.13. Comparison of potential causative variants identified with WES using the MARS subsequent to PSGT in the index patient and her mother diagnosed with breast cancer

Chr: pos	Rs number	Gene name	Ref/ Alt allele	Patients	Index patient Genotype	Mother Genotype	Index patient Coverage	Mother Coverage	Variant function
1:144873963	.	PDE4DIP	T,C	Both	Wildtype	Wildtype	1010	38	Missense
1:144917827	rs11295415	PDE4DIP	CA,C	Both	Hetero	Hetero	291	63	frameshift-deletion
1:144923823	rs11268079	PDE4DIP	**,T	Both	Homo	Homo	503	47	Splicing
1:155160668	.	MUC1	G,C	Daughter	Hetero		248		Missense
1:179077247	.	ABL2	G,A	Daughter	Hetero		144		Missense
3:142224045	.	ATR	CTTT,C	Daughter	Hetero		161		nonframeshift-deletion
4:62812800	.	LPHN3	A,T	Daughter	Hetero		120		Missense
5:112176324	.	APC	GA,G	Daughter	Hetero		300		frameshift-deletion
5:131924480	rs139372231	RAD50	C,T	Both	Hetero	Hetero	359	16	Missense
6:94120780	.	EPHA	A,G	Daughter	Wildtype		194		Missense
6:160471661	.	IGF2R	C,G	Daughter	Hetero		574		Missense
6:160471662	.	IGF2R	T,A	Daughter	Hetero		573		Missense
7:91652178	rs10644111	AKAP9	A,AAAC	Daughter	Hetero		157		nonframeshift-insertion
7:126086302	.	GRM8	C,T	Daughter	Hetero		240		Missense
12:49437499	rs186948725	KMT2D	G,A	Daughter	Hetero		337		Missense
12:49444818	.	KMT2D	A,G	Daughter	Wildtype		380		missense
20:41514485	.	PTPRT	G,C	Daughter	Hetero		222		missense
22:36717864	.	MYH9	C,T	Both	Hetero	Hetero	593	80	synonymous, splicing
x:70586328	.	TAF1	A,G	Daughter	Homo		106		missense

**TAGGAAGAAACGGCA

Due to the important associations found with SNPs involved in the folate-methylation pathway in phase II of this study, we shifting our focus to common variants implicated in the DNA mismatch repair that could serve as potential pharmacogenetics treatment targets. Two functional polymorphisms in the methionine synthase (MTR) and methionine synthase reductase (MTRR) genes included in the Comprehensive Cancer Panel, previously shown to be involved in the folate-homocysteine-methylation pathway, were identified in the three breast cancer patients (Table 5.14).

Table 5.14. Comparison of genotypes resulting from WES (aligned to MARS) selected for investigation and validation by Sanger sequencing/ Taqman genotyping in the three breast cancer patients included in this study. Amino acid positions are noted according to ExAC Browser and PROVEAN.

rs number/ chr pos.	Gene	WES Genotypes (a/b)	WES Genotypes (c)	Variant effect	Amino acid change	Sanger sequencing/ TaqMan genotyping	PROVEAN/ SIFT prediction
Rare variants							
rs139372231	RAD50	Hetero/ Hetero	Wild type	Missense	R385C	Confirmed	Neutral/ damaging
22:36717864	MYH9	Hetero/ Hetero	Wild type	Splice region	G236G	Confirmed	Neutral/ tolerant
1:155160668	MUC1	Hetero/ --	Wild type	Missense	Q67E	Confirmed	Probably damaging/ Deleterious
6:160471661	IGF2R	Hetero/ --	Wild type	Missense	L981V	Not confirmed	Neutral/ damaging
6:160471662	IGF2R	Hetero/ --	Wild type	Missense	L891H	Not confirmed	Deleterious/ damaging
Common variants							
rs1130409	APEX1	Homo TT/ Homo TT	Homo GG	Missense	D148E	Confirmed	Neutral/ tolerant
rs1805087	MTR	Wild type/ Hetero	Hetero	Missense	D919G	Confirmed	No record
rs1081394	MTRR	Hetero/ Hetero	Homo	Missense	I22M	Confirmed	No record

a= index patient, b= mother, c= unrelated breast cancer patient

MTR 2756 A>C (rs1805087) (MAF=0.22) was detected with the use of both the MARS and hg19 whereas MTRR 66 A>G (rs1081394) (MAF=0.36) was detected with hg19 and not with the MARS. A more frequently occurring SNP in the Apurinic endonuclease 1 (APEX1) gene (rs1130409) (MAF=0.37) in the family and not in the unrelated breast cancer patient. This variant was detected with use of the MARS but not with alignment of the sequence reads to the hg19 reference genome.

5.5.4. Verification of gene variants identified using WES

WES identified an intermediate-risk variant (frequency <0.001) in the RAD50 gene (rs139372231), shared by both the index patient and her daughter, and a rare variant in the MUC1 gene (Q67E). Heterozygosity for RAD50 rs139372231 and MUC1 Q67E detected in the family was confirmed by allele-specific Taqman genotyping. Although this variant was absent in an additional 456 individuals (160 patients and 296 controls) selected from a genomics database resource, one control individual (36 year female) with the same rare variant was identified using TaqMan genotyping.

Extended family screening for the same variant excluded the RAD50 variant in a cousin diagnosed with breast cancer at the age of 43 years. Figure 5.15 illustrates the clinical, histopathological and genetic heterogeneity in the family of the index patient.

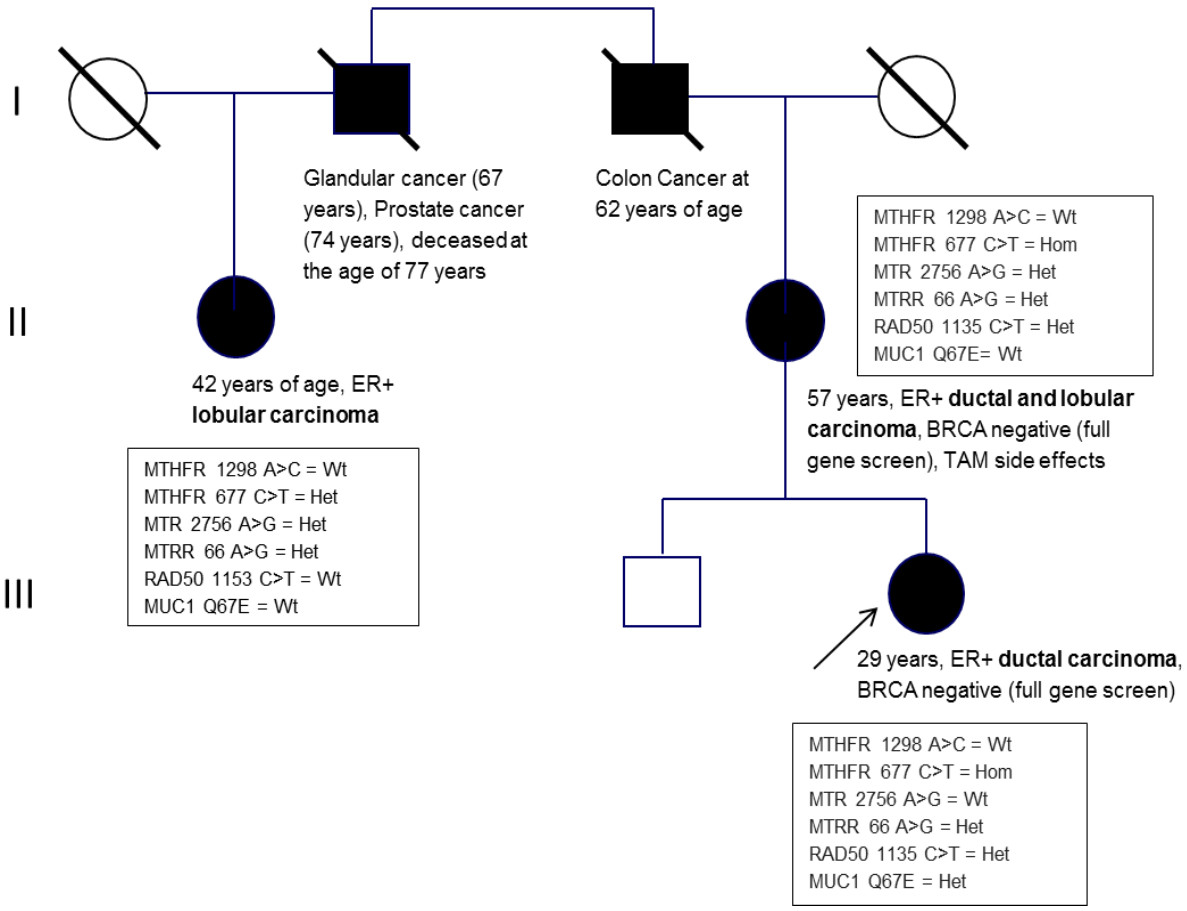


Figure 5.18. Pedigree of family selected for exome sequencing, depicting the clinical (age of onset), pathology (cancer subtypes) and genetic (gene variation including 2 SNPs from the CVD multi-gene assay) heterogeneity in the family. Wt = homozygous for the common allele, Het = heterozygous, Hom = homozygous for the rare allele.

Assessment of allelic combinations of MTHFR 677 C>T, MTHFR 1298 A>C, MTR 2756 A>G and MTRR 66 A>G revealed that the TAGG combination is significantly associated with ER-positive breast cancer (p=0.034), after adjustment for age, ethnicity and BMI (Table 5.15). The effect of MTHFR 677 C>T, MTHFR 1298 A>C, MTR 2756 A>G and MTRR 66 A>G TAGG allelic combinations are depicted in a barplot which compares breast cancer patients stratified by ER status with respect possibly carrying and not carrying the TAGG genotype (Figure 5.15).

Table 5.15. P-values for genotype distribution and additive allelic differences in ER status.

Genotype	Genotype		Additive allelic	
	Unadjusted	Ethnicity-adjusted	Unadjusted	Ethnicity-adjusted
MTHFR 677 C>T	0.644	0.747	0.570	0.904
MTHFR 1298 A>C	0.529	0.567	0.892	0.987
MTR 2756 A>G	0.259	0.333	0.348	0.391
MTRR 66 A>G	0.524	0.530	0.380	0.318
4 SNPs CAAA			0.273	0.205
4 SNPs CAAG			0.236	0.304
4 SNPs CAGA			0.416	0.506
4 SNPs CAGG			0.554	0.638
4 SNPs CCAA			0.379	0.544
4 SNPs CCAG			0.807	0.828
4 SNPs CCGA			0.719	0.702
4 SNPs CCGG			0.122	0.150
4 SNPs TAAA			0.930	0.801
4 SNPs TAAG			0.854	0.525
4 SNPs TAGA			0.053	0.095
4 SNPs TAGG			**0.019**	**0.034**
4 SNPs TCAA			0.163	0.165
4 SNPs TCAG			0.174	0.132
4 SNPs TCGG			0.333	0.435

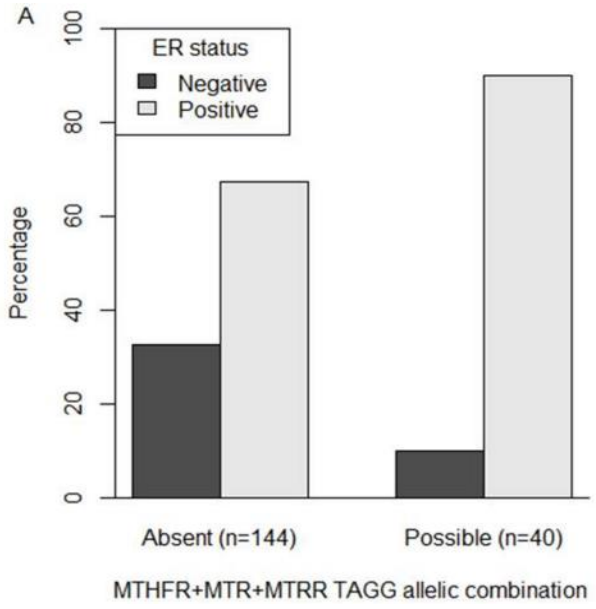


Figure 5.19. Effect of MTHFR 677 C>T, MTHFR 1298 A>C, MTR 2756 A>G and MTRR 66 A>G TAGG allelic combinations. Barplot comparing ER-negative and ER-positive breast cancer patients with respect to not carrying and possibly carrying the TAGG genotype.

5.6. PHASE III DISCUSSION

The aim of the study was to identify potential causative gene variants in three genetically uncharacterised breast cancer patients from two unrelated families in a comparative study of the publicly-available hg19 and the ethnically concordant MARS. The newly-developed EPA for breast cancer was used to identify genetically uncharacterised patients for extended mutation analysis after exclusion of genetic risk factors that could explain or contribute significantly to the disease profile. These include obesity in patients without a family history of breast cancer and high-penetrance BRCA mutations in patients with early-onset breast cancer, in order to identify the cause for breast cancer in the selected genetically uncharacterised patients. Confirmation of presence or absence of gene variants included in the EPA contributed to the analytical validation of the sequencing platform used for WES in this study, as confirmed with use of both hg19 and the MARS.

5.6.1. Discordance between exome variant calling using different reference genomes

Variant calling using both hg19 and the MARS following WES performed in DNA samples of the three patients and three population-matched controls, revealed >20% discrepancy in the number of gene variants identified in the same samples. Alignment of sequence reads to hg19 resulted in a substantially higher number of variants detected with the exomes of the study participants, most likely false-positives due to the presence of minor alleles in hg19. This incorrect designation of minor alleles may adversely affect the sequence read alignment as well as subsequent variant calling used in WES. The result is that in individuals who possess two copies of the minor allele contained in the reference genome, the gene variant at that locus will not be detected by variant calling and the opportunity for interpretation of clinical relevance will therefore be missed. Homozygotes for incorrectly designated minor alleles would furthermore be reported as potential disease-associated variants at such loci. This highlights the importance and supports the development of a catalog of clinically annotated variants (SNP knowledge database) (Kotze et al. 2014) as a vital component of the PSGT platform containing information extracted from the literature on the risk allele and disease associations previously replicated in different populations (Kotze et al. 2015). A relatively new resource released in 2004 is The Catalogue of Somatic Mutations in Cancer (COSMIC) which is currently the most comprehensive global resource on somatic mutations in human cancers. Full curation of the scientific literature for genes known to be involved in cancer, as defined in the Cancer Gene Census (Futreal et al. 2004), improved the clinical usefulness of published somatic mutation data and has been further extended to include tumour resequencing results from the Cancer Genome Project (CGP) at the Wellcome Trust Sanger Institute. Results from meta-analyses are also freely available and accessible online from databases including the Single Nucleotide

Polymorphism Database of the National Centre for Biotechnology Information (www.ncbi.nlm.nih.gov/SNP), SNPedia that gathers information from PubMed on a daily basis to support genome annotation and interpretation (http://www.snpedia.com/index.php/SNPedia) and ExAC Browser (Beta)/ Genome Aggregation Consortium (http://exac.broadinstitute.org/) who aggregate and match exome sequencing data from various large-scale sequencing projects including 1000 Genomes, The Cancer Genome Atlas (TCGA) and many more.

5.6.2. BRCA1 and BRCA2 mutation screening

BRCA1 plays a pivotal role in maintaining genomic stability and acts as a tumour suppressor. It is a versatile protein that interacts with DNA damage sensing proteins, DNA damage response effectors and cell cycle regulators. It therefore functions in multiple DNA repair pathways including homologous recombination, single-strand annealing and non-homologous end joining as well as in cell cycle checkpoint regulation (Dever et al. 2012; Roy et al. 2012; Medema and Macůrek 2012). Numerous BRCA1 and BRCA2 polymorphisms have been reported to confer small increased risk to various types of cancer. A BRCA2 SNP (rs144848) identified in the index patient was reported to confer a 1.31x increased risk (CI: 1.07-1.61) for breast cancer (Healey et al. 2000). Furthermore, a BRCA1 SNP (rs799917) identified in this index patient was previously shown to impact the interaction between BRCA1 messenger RNA (mRNA) and miR-638 (Nicoloso et al. 2010) and has recently been associated with increased risk of esophageal cancer (Zhang et al. 2013). This SNP as well as two other BRCA1 SNPS (rs16941 and rs1799966) also identified in the unrelated TNBC patient without a positive family history were shown to interact with environmental factors (obesity, smoking, alcohol) to increase risk for esophageal cancer (Nicoloso et al. 2010) and postmenopausal breast cancer (Figueiredo et al. 2011; Ricks-Santi et al. 2013). WES showed 100% concordance with the results of a BRCA1/2 full gene screen performed in the index patient, suggesting that this relatively new technology may be sufficient in detecting variants in the BRCA1/2 genes. Given the increasing availability and decreasing cost of NGS, this technology may in the very near future replace limited BRCA1 and BRCA2 mutation screening in eligible patients, to provide a valuable healthcare resource for disease diagnosis and characterization over the course of an individual's lifespan. Yeo et al. (2014) compared the performance of the Ion Torrent platform using different versions of the Torrent Suite (TS) software for detection of indels in the BRCA1 and 2 genes in 17 breast cancer patients. They contended that although older software versions have generated many false-positive calls in detecting insertions and deletions which potentially hindered its clinical utility, new versions of the TS have shown improvements in indel calling sensitivity and specificity. Despite The Genome Analysis Toolkit (GATK) and The Sequence Alignment/Map format (SAM)-tools's superior sensitivity over the TS, proper computational indel calling analysis

is able to maximize the sensitivity and specificity at the single base level, which will enable future application of this technology in clinical genetic testing.

15.6.3. Comprehensive cancer panel screen in familial breast cancer

Variable penetrance of breast cancer as demonstrated by Kotze et al. (2005) in a relatively large South African family with the founder BRCA1 mutation E881X supports the view that other genetic and/or environmental factors may modify breast cancer risk not only in the general population, but also in familial cases (King et al. 2003). High penetrance breast cancer susceptibility genes account for up to 50% of familial breast cancer risk (Miki et al. 1994; Zhong et al. 2013). RAD50 has been documented as an intermediate-risk gene for breast cancer susceptibility allele, attributable to its critical role in the aberrant repair of DNA double strand breaks (Damiola et al. 2014).

The RAD50 variant (rs139372231) is a C>T base change detected at nucleotide position 1153 that results in an amino acid change from Arginine to Cysteine at position 385. The RAD50 protein forms a complex with the MRE11 and NBN proteins and binds to strands of damaged DNA, holding broken ends together during the repair process. The RAD50/MRE11A/NBN protein complex work together to mend damaged strands of DNA, thereby preventing the accumulation of DNA damage that may trigger uncontrollable cellular division. This complex therefore aids maintenance of genetic information stability (Heikkinen et al. 2006). In addition, the RAD50/MRE11A/NBN complex interacts with the protein encoded by the ataxia telangiectasia mutated ATM gene, which plays an essential role in recognizing broken strands of DNA and coordinating their repair (Deshpande et al. 2014). Kinoshita et al. (2015) suggested that MRE11 is important in maintaining the structural arrangement of RAD50 in this protein complex. Mutations in the RAD50 gene lead to the production of an abnormally small, non-functional or missing RAD50 protein which results in the inability to respond effectively to DNA damage (Bartkova et al. 2008). Dai et al. (2014) demonstrated that RAD50 and inositol polyphosphate 4-phosphatase B (INPP4B) expression levels have a synergistic influence on breast cancer survival, possibly through their effects on treatment response. INPP4B has been identified as a tumour suppressor mutated in various epithelial cancers by inhibiting PI3K/Akt signaling. However recently, tumour-promoting features of INPP4B in leukemia and breast cancer have been uncovered. Due to the combined effect reported between INPP4B and various DNA repair genes including RAD50 (Ip et al. 2015; Kofuji et al. 2015; Dai et al. 2015), we have investigated and excluded co-inheritance of deleterious mutations in the family.

RAD50 rs139372231 was scored as being benign by PolyPhen-2 and PROVEAN. In contrast, this variant was predicted to have a potentially damaging effect by the SIFT algorithm. RAD50

rs139372231 detected in our exome-sequenced family has previously only been found in African, South Asian and European (Non-Finnish) populations (http://exac.broadinstitute.org/). Other, more common RAD50 gene variants have however been detected in Swedish, Norwegian, Finnish and Icelandic populations (Heikkinen et al. 2006; Heikkinen et al. 2003). To date, a splicing mutation in the RAD50 gene (IVS3-1G>A) leading to a translational frameshift as well as a deletion Finnish founder mutation (RAD50 687delT) not present in the other Nordic cohorts was identified by Heikkinen et al. (2003). In this study, the authors concluded that their findings support an effect for RAD50 and NBS1 haploinsufficiency on genomic integrity and susceptibility to breast as well as ovarian cancer. Tommiska et al. (2006) screened 435 European and 46 Finnish familial breast cancer cases and concluded that RAD50 mutations are rare in familial breast cancer and contribute a negligible amount to overall genetic risk for cancer. The authors contended that RAD50 only provides a very minor contribution to familial breast cancer predisposition in UK and Finland. Dai et al. (2015) found that RAD50 and INPP4B expression levels have a synergistic influence on breast cancer survival, possibly through their effects on treatment response. Abuzeid et al. (2009) suggested that the use of targeted RAD50 disruption could be a novel chemo-sensitizing approach for cancer therapy in the context of chemo-resistance. This strategy could be applicable to several types of cancers that demonstrate chemo-resistance and may exert a positively influence on breast cancer treatment leading to improved therapeutic outcomes. Heterozygosity for RAD50 rs13937223 identified in the index patient and her mother was verified by IGV and TaqMan genotyping of their DNA samples. Although this variant was detected in one of 456 unrelated individuals screened, and shown to be absent in a cousin of the index patient' mother, it may express a deleterious effect against a background of risk conferred by variation in the MTHFR, MTR and MTRR genes shown to be associated with ER status (phase II of this study). Functional polymorphisms in these genes underlying dysfunction of the folate-methylation pathway confers risk at the population level (Choi and Mason 2002; Macis et al. 2007; Maruti et al. 2009; Peerbooms et al. 2011; Wu et al. 2013; Weiwei et al. 2014; Kakkoura et al. 2015).

The difference in tumour subtypes between the index case and her mother - albeit the same MTHFR rs1801133 genotype associated with the luminal B subtype (Naushad et al. 2011) - prompted the search for gene variants which may be responsible for their discrepant HER2 statuses. Two HER2 variants previously studied in relation to breast (rs1136201) (AbdRaboh et al. 2013) and endometrial cancer risk (rs1810132) (Tong et al. 2009) were identified in the index patient but not in her mother, nor the unrelated breast cancer patient. Three variants in the MUC1 and IGF2R genes previously associated with HER2 were identified in the daughter but not in her mother, as was confirmed with the use of Sanger sequencing for MUC1. MUC1 is a heterodimeric oncoprotein that associates with HER2 at the surface of breast cancer cells (Li et

al. 2003; Kufe et al. 2013). This gene has been documented as a moderate penetrance gene in the development of breast cancer (Siroy et al. 2013; Raina et al. 2014). Raina et al. (2014) suggested that, in carcinoma cells with sustained activation of the epithelial-mesenchymal transition, the MUC1-C protein is positioned to interact with the HER2 complex and promote activation of the HER2 pathway in HER2-overexpressing breast cancer cells, thereby promoting their growth and clonogenic survival. Rakha et al. (2005) assessed the prognostic value of several mucins in 1447 cases of invasive breast carcinoma. They concluded that most breast cancers express MUC1, MUC3 and MUC4 and that MUC1 and MUC3 in particular are potential prognostic indicators, with MUC1 having the strongest relationship with patient outcome. Kufe and Kharbanda (2014) have shown that MUC1 peptide therapy can be used successfully in conjunction with an anti-HER2 agent such as trastuzumab to render resistant cells sensitive to treatment. Kufe and Kharbanda (2014) have shown that MUCl peptide therapy can be used successfully in conjunction with anti-HER2 trastuzamab to render resistant cells sensitive to treatment. Moreover, Mountzios et al. (2014) evaluated the prognostic significance of components of the insulin-like growth factor (IGFR) pathway across different breast cancer subtypes and found that aberrant expression of important components of the IGF1R-alpha/IGF2R complex are associated with better clinical outcomes in patients with hormone-receptor positive, HER2-negative, node-positive early-onset breast cancer. They concluded that their results support the important interplay between the IGFR pathway and hormone receptors and suggest their potential role as molecular targets for therapeutic intervention in hormone-receptor positive disease.

The genetic heterogeneity detected in the family is compatible with a disturbance in the folate pathway which has been documented as one of the three most important mechanisms in the pathogenesis of cancer (Wishart 2015). The presence of RAD50 rs139372231 in combination with MTHFR rs1801133 as a BRCA-modifier gene (Pepe et al. 2007) as well as other risk-contributing SNPs is consistent with an intermediate risk profile as reflected by the family history. According to several authors (Antoniou et al. 2008; Pharoah et al. 2008, Ricks-Santi et al. 2013), the overall importance of polymorphisms remains high due to their effect which appears to be multiplicative, such that individuals possessing several polymorphisms may have a significantly increased risk of breast cancer. Gene-gene interaction between RAD50 rs139372231 and other gene variants implicated in a dysfunctional folate pathway such as MTHFR rs1801133, methionine synthase MTR (rs1805087) and methionine synthase reductase (MTRR) (rs1081394) may, as Bodmer and Tomlinson (2010) suggested, be sufficient in accounting for the breast cancer in the index case and her mother. Extended mutation screening of the mother's cousin diagnosed with breast cancer at the age of 53 years, excluded the presence of RAD50 rs139372231, although she was heterozygous for MTHFR 677 C>T,

MTR 2756 A>G and MTRR 66 A>G. In the unrelated TNBC patient without a family history of cancer, neither the RAD50 R385C nor MTHFR rs1801133 were identified.

5.6.4. Analysis of low penetrance variants implicated in DNA mismatch repair

Cancer development may involve three types of gene families: tumour suppressor genes such as BRCA1/2, proto-oncogenes including HER2 routinely assessed in all breast cancer patients, and stability genes responsible for DNA repair. When stability genes such as MTHFR are affected by functional polymorphisms, mutations may occur at significantly higher rates in other genes involved in a wide spectrum of diseases. MTHFR, MTR and MTRR are involved in the folate pathway essential for DNA repair and methylation. Functional variants in the MTHFR gene resulting in reduced enzyme activity have been shown to cause alterations in chromosomal recombination and abnormal chromosome segregation (Friso et al. 2002). It has been found in many studies that common SNPs in the MTHFR gene are involved in chromosomal nondisjunction and linked to an increasing risk of trisomy 21 (Hobbs et al. 2000), as well as other chromosomal aneuploidy (Hassold et al. 2001, Oliveira et al. 2008). The OncoScan™ microarray for analysis of copy number and somatic mutations in cancer could be used to confirm the diagnosis of malignant melanoma, stratify the risk of recurrence and direct pharmacologic interventions (Chandler et al. 2012). Sangrajrang et al. (2010) reported an association between MTR and MTRR and breast cancer risk in a study performed in 1067 individuals. Variation in these genes may therefore disrupt folate homeostasis and confer significant risk of breast cancer, attributable to impaired DNA repair associated with the epigenetic process of DNA methylation (Macis et al. 2007; Weiwei et al. 2014). To avoid the accumulation of DNA damage induced by environmental stimulants (e.g. alcohol, smoking, chemical agents and high dietary fat intake), cells have developed precise and effective DNA repair systems (Zhang et al. 2011). Base excision repair (BER) is the major DNA repair pathway (Wilson et al. 2011) which recognizes and repairs base modifications and single strand breaks (SSB) (Maynard et al. 2009). BER involves four steps including (1) base lesion recognition, excision, and cleavage of an aberrant site; (2) terminal-processing of SSB ends to generate 3’ hydroxyl (OH)/5’ phosphate group ends; (3) gap-filling after lesion excision; and finally (4) nick sealing by DNA ligases (Hegde et al. 2012).

APEX1 is one of three key enzymes responsible for the incision of the apurinic/apyrimidinic sites and the generation of 3’-OH termini (Tell et al. 2010). APEX1 is a multi-functional enzyme that exhibits DNA repair activity and plays a role in the reductive activation of many transcription factors which are involved in various cellular processes such as cell survival, growth signalling and inflammatory pathways (Gaiddon et al. 1999; Bapat et a. 2008; Tell et al. 2010). According

to dbSNP in NCBI, APEX1 rs1130409 results in a T>G change and an amino acid change from Asp to Glu at position 148. However, studies of the prevalence of this variant in different populations indicate that the minor allele in Caucasian is in fact the T-allele and not the G-allele as in other population groups. Exome sequence-alignment to the ethnically concordant MARS correctly unmasked this variant in the Caucasian family where both mother and daughter are homozygous for the T-allele, whereas it was missed by the ethnically ignorant hg19. Although a very commonly occurring polymorphism (MAF=0.36), and despite the fact function prediction using bioinformatics tools was negative, the deleterious effect of the APEX1 variant has been proven by an *in vitro* study performed by Hu et al. (2001). In addition, APEX1 has been documented as an emerging therapeutic target across the disease spectrum. APEX1/redox effector factor-1 (Ref-1) repair and redox functions are activated as a part of the DNA repair response to maintain cellular genomic integrity as well as regulate various transcription factors in cell-survival pathways. Natural plant products and phytochemicals (including soy isoflavones, curcumin, resveratrol, tea polyphenols and decursin) potentially modulate the expression and the repair/redox activities of APE1/Ref-1, thus representing a possible therapeutic intervention against various human diseases associated with reactive oxygen/nitrogen species (ROS/RNS)-mediated oxidative stress, including cancer, cardiovascular and neurodegenerative disorders (Thakur et al. 2014). Pharmacogenetic studies have traditionally focused on genes involved in processes such as drug metabolism however attention is increasingly shifting to the effects of genetic variations in drug targets and associated pathway components on drug responses (Krejsa et al. 2006).

5.6.5. Verification of gene variants identified using WES

APEX1 rs1130409 was identified in the family using the MARS but not hg19, representing a locus of conflicting assignment of the major/minor allele using the two different reference genomes. In the dbSNP database, APEX1 rs1130409 is documented as a T>G change while studies performed in various population groups document that the T-allele is in fact the minor allele in Caucasians. Based on this finding it is clear that the major/minor allele status is population-dependent, emphasising the importance of using an ethnically concordant reference genome in WES. Interestingly, MTRR rs1081394 was detected using hg19 but not with the MARS in all three breast cancer patients studied. Differences in detection of this SNP possibly occurred due to differences in allocation of the minor/risk allele according to the MAF as previously determined in world populations. This discrepancy warranted verification by Sanger sequencing/ Taqman genotyping, which then confirmed the results obtained using hg19.

Subsequent to the identification of a large number of false-positives and false-negatives caused by incorrectly designated minor alleles by hg19 vs MARS, a list of potentially causative variants identified by WES filtered on a MAF of less than 1% and subsequently possible low-penetrance pharmacogenetics targets were prioritized for verification in the Pathology Research Facility (PRF) laboratory using Sanger sequencing. Except for MTRR rs1081394 due to the different allele assignment in hg19 vs MARS, sequencing results of all of the gene variants investigated in this study corresponded to WES genotype calls using the MARS. This high level of correspondence increased our confidence in using the MARS as a standard reference genome when applying WES in uncharacterized patients.

The assessment of the aforementioned non-deterministic risk modifier variants across the diagnostic spectrum as part of a multi-gene assay has led to extended analysis in genetically uncharacterised patients using the PSGT approach (Kotze et al. 2015). Homozygosity for the MTHFR rs1801133 variant in both mother and daughter with early-onset breast cancer raised the question of whether or not this genotype is sufficient to confer significant disease risk in a synergistic/additive manner together with other intermediary risk alleles. The involvement of MTHFR rs1801133 in the development of breast cancer and associated comorbidities including depression (Delport et al. 2014), CVD and treatment-related side effects is in keeping with the role of MTHFR as a target for both prevention and treatment across the diagnostic boundaries. Although genetic variants identified by the EPA, of which some were excluded prior to WES, do not confer significant individual risk for breast cancer, cumulative data supported by WES suggests that co-inheritance of multiple low-penetrance mutations (in genes such as MTHFR, MTR, MTRR, APEX1) and rare variants (eg. in RAD50 gene) may explain the majority of BRCA mutation-unrelated familial disease (Antoniou and Easton 2006; Gracia-Aznarez et al. 2013).

In this exploratory study, we built upon our prior experience with prospective molecular profiling (Grant et al. 2013, 2015). This experience has taught us that any new method or technology, however unlimited the potential in its respective capacity, is subject to flaws that can only be rectified through on-going discovery and frequently the process of trial and error. Since NGS is an emerging technology, there were expected and unforeseen challenges encountered during the course of the study. Foreseen challenges included the discordance expected between the use of two different reference genomes hg19 and the MARS, resulting in the detection of an unprecedented amount of false-positives and false-negatives. An unforeseen challenge realized through the course of the study was the flawed processing of variant call files which can only be detected with thorough evaluation of resulting variants against SNP/mutation databases, prior to the commencement of Sanger sequencing and genotyping for verification. Even if there are no technical or logical delays in NGS, the interpretation and application of this information in clinical

practice as confirmed in this study, is fraught with uncertainty. Not all gene variants identified by WES will have a deleterious effect on protein function, so the mere identification of variants - even when validated by Sanger sequencing - does not ascertain or guarantee clinical usefulness in relation to risk prediction or disease management.

There are considerable challenges associated with the discovery and testing of rare predisposition alleles, many of which are illustrated by the issues associated with variants of unknown significance in the Mendelian cancer predisposition genes including BRCA1 and BRCA2. Bodmer and Tomlinson (2010) reckoned that whether or not rare variants end up filling the heritability gap, it is imperative to look for them alongside common variants. This study has demonstrated the feasibility of WES to achieve this level of comprehensive genomic testing. The detection of RAD50 rs139372231 in two affected members of the index family while the niece / cousin with ILC tested negative, and detection in only one control (out of a total of 457 individuals screened) confirms that it is a rare variant and suggests a possible role in combination with other functional variants in the development of breast cancer in the studied family.

Several novel variants were detected in genes described above, highlighting the genetic differences in South African Caucasian women and emphasising the fact that complete analysis of the genes, rather than looking for a few known variants, may be required to identify these high risk women. At loci where the major alleles of variants with potential high impact are absent from hg19, potential causative variants may be missed. Our findings highlight the fact that the presence of minor alleles relating to disease-susceptibility in hg19 should be taken into account during data interpretation for assessment of personal disease risk, since it is dependent on the identification and functional annotation of genetic variants that may be discordant or erroneous. In this study, the need for improvement of hg19 used as an intermediate bioinformatics step for data translation between the laboratory and clinic was addressed by the use of a MARS, which effectively reduced the number of false variants identified.

Given that ~85% of disease-associated variants reside within exons, the remainder of variants in non-exonic (intronic, UTR, intergenic, etc.) regions may be missed by WES. WES however has the potential to accelerate the discovery of new susceptibility genes which may provide a means to identify novel therapeutic targets to optimize patient care based partly on the assessment of genetic risk factors. Nevertheless, this era of vast data generation and identification of novel genes, while opening new roads in cancer susceptibility, still needs to be treated with caution, as the gaps in knowledge highlighted in this study have indisputable implications for data translation in the clinical setting. Genetic counselling for novel or unexpected findings may present a challenge that needs to be addressed using a

multidisciplinary approach which incorporates all relevant information in relation to current and future risk in each individual. Our findings support previous WES results indicating that the majority of genetically uncharacterised breast cancer may be caused by a combination of low-penetrance mutations exerting their effect against a high-risk clinical background. WES applied beyond the limited scope of single-gene testing has the potential to detect both known and novel mutations across diagnostic boundaries to facilitate prevention of cumulative risk underlying the development of breast cancer and associated comorbidities.

The genomic database utilized in this study proved to be a valuable resource for validation of variants detected by WES as well as the interpretation of those of uncertain clinical significance. Its on-going development provides an important resource for health outcomes studies and allows for follow-up and careful monitoring of patients over time. Ultimately, it is envisioned that the patient database enabling extensive local research will continue to inform clinical decision making and optimize delivery and implementation of a comprehensive and truly individualized genomics-based healthcare service for breast cancer patients based on disease patterns relevant to the South African population.

CHAPTER 6

CONCLUSIONS

This study contributed to the development of a pathology-supported genetic testing (PSGT) framework for whole exome sequencing (WES) beyond the limitations of single-gene BRCA testing in South African breast cancer patients. The cardiovascular disease (CVD) multi-gene assay that combines well-characterised clinically actionable single nucleotide polymorphisms (SNPs) in key overlapping disease pathways for multifactorial diseases in a way similar to founder mutations for single high-penetrance genes such as BRCA1/2, was used as part of a newly developed exome pre-screen algorithm (EPA) to identify genetically uncharacterised patients or non-responders to treatment as eligible for WES (Figure 6.1).

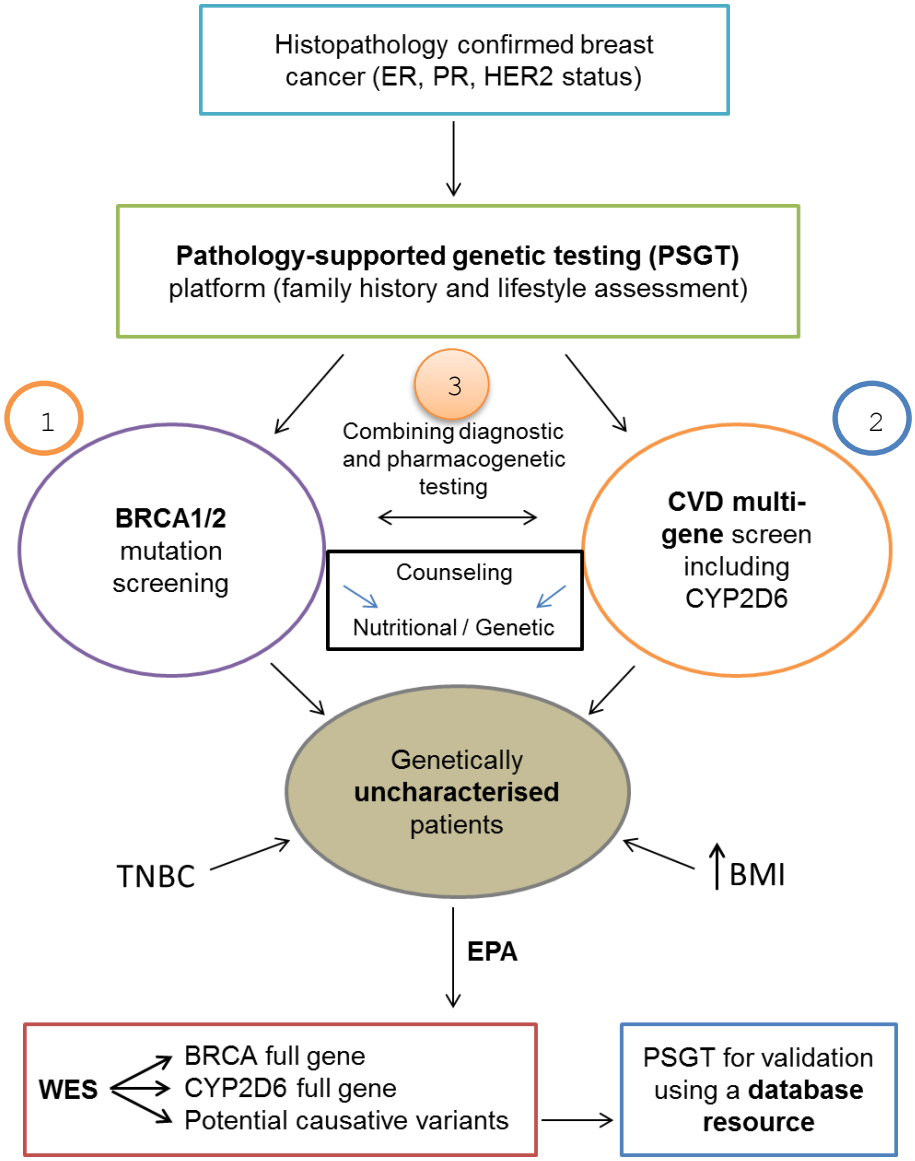


Figure 6.1. The PSGT platform used to develop an exome pre-screen algorithm (EPA) for selection of genetically uncharacterised patients for WES. 1) Diagnostic BRCA testing is offered as a routine service according to standard referral guidelines (aged <40 years, bilateral breast cancer, additional ovarian cancer, one 1st degree relative with breast or ovarian cancer, > 2nd

degree affected relatives and male breast cancer in a relative) or 2) the chronic disease risk screen is offered to patients receiving hormone therapy, are at high risk for tumour recurrence (family history, BRCA1/2 positive), or are required to take potentially competing antidepressants or 3) combined diagnostic and pharmacogenetics testing to explain the presence of comorbidities or predict drug response/recurrence risk. Where extended mutation analysis of the entire BRCA1 and 2 genes as well as the CYP2D6 gene were unable to explain breast cancer or the occurrence of drug side effects/failure, WES is performed to identify potential novel causative genes/mutations.

The lack of a clear road map to follow in navigating and interpreting vast amounts of genomic information generated by WES impedes the rapid translation of genomic discoveries into practical benefits for patients and their at-risk family members. PSGT is positioned to bridge the gap between massive amounts of genomic information generated by advanced technologies in order to achieve a high degree of concordance between genotype and phenotype. This study was performed in three phases, starting with single-gene CYP2D6 genotyping extended to a multi-gene panel for identification of cardio-metabolic risk factors as part of a screening step to identify genetically uncharacterized breast cancer patients eligible for WES.

By combining diagnostic and pharmacogenetic testing of breast cancer patients using the PSGT approach, the occurrence of breast cancer recurrence and associated co-morbidities including depression, cognitive impairment and deep vein thrombosis could at least in part be explained by genotype. The age at diagnosis/onset of breast cancer was significantly lower in ER-negative than ER-positive patients after adjustment for ethnicity ($p=0.022$), while the presence of the relatively common multi-functional E4 allele of the APOE polymorphism was shown to reduce the mean age of disease onset by an average of 10 years regardless of ethnicity, alcohol consumption, current smoking or family history of cancer/BRCA mutation status ($p=0.003$). The APOE allelic effect was 9% (95% CI: 4 to 15). BMI was found to be significantly higher in ER-positive compared to ER-negative breast cancer patients after adjustment for age, ethnicity, and family history of cancer/BRCA mutation status ($p=0.035$). In addition, BMI showed a positive allelic association with the relatively common low penetrance MTHFR 1298 A>C mutation ($p=0.01$). Presence of the risk-associated C-allele, previously linked to sporadic breast cancer, was associated with an 8% (95% CI: 3 to 14) increase in BMI. The finding of a significant positive association between BMI and the fat score ($p=0.032$) in the implementation data set as assessed by the medical and lifestyle questionnaire completed for this subgroup, highlights the fact that dietary adjustments guided partly from the genetic background may not only prevent obesity, but also a substantial proportion of breast cancer in the general population.

Based on the above findings, CYP2D6 pharmacogenetic testing, as well as APOE and MTHFR genotyping performed as part of an extended CVD multi-gene assay, were incorporated into the EPA used in this study to select genetically uncharacterized South African patients for extended NGS. As part of this PSGT approach, each individual was asked to complete a medical and lifestyle questionnaire required to interpret the genetic results in a clinical context. After evaluation of the EPA in each individual considered for extended genetic testing, three patients were selected for WES performed as part of this study: 1) two ER-positive breast cancer patients from the same family without mutations in the BRCA1/2 genes (familial breast cancer) and 2) one triple-negative breast cancer patient (TNBC) not previously screened for BRCA mutations due to absence of a family history of cancer (sporadic breast cancer). In accordance with the EPA, CYP2D6 allele 4, the APOE E4 allele and MTHFR 1298 A>C were excluded as possible contributory factors to disease development or medication side effects prior to WES performed in these patients. Due to the significantly higher BMI observed in the ER-positive compared with ER-negative South African breast cancer patients studied, obesity and other lifestyle risk factors were also identified as important criteria evaluated as part of the EPA. In the patient with TNBC diagnosed with this severe disease subtype at the relatively young age of 43 years, the BMI was within the normal range (20.4 kg/m2). In the family with inherited breast cancer, both the index patient (BMI of 27.5 kg/m2) and her mother (BMI 27.8 kg/m2) were overweight. Both breast cancer patients from the same family were homozygous for the MTHFR 677 C>T mutation assessed as part of the EPA, while the TNBC patient tested negative for this low-penetrance mutation. Previous studies reported better progression-free survival in breast cancer patients with MTHFR 677 C>T found to be linked to the luminal-type breast cancer. This is in accordance with results obtained from a 70-gene microarray profile (MammaPrint) performed in 2007 on both tumours of the mother presenting with bilateral breast cancer. This patient has remained disease-free over the past 8 years despite termination of tamoxifen after one year of treatment due to severe side effects. WES was considered important in this family due to the fact that homozygosity for MTHFR 677 C>T in both the mother and daughter may not be sufficient to cause breast cancer in the family. Notably, human epidermal growth factor receptor-2 (HER2) status was negative in the mother, while her daughter's tumour was HER2-positive.

The next step involved variant calling using both hg19 and the MARS following WES performed in DNA samples of the three patients and three population-matched controls, which revealed >20% discrepancy in the number of gene variants identified in the same samples. After exclusion of a large number of false-positives caused by minor alleles in hg19, a list of potentially causative variants identified by WES filtered on a minor allele frequency (MAF) of less than 1% were prioritized for verification in the Pathology Research Facility (PRF) laboratory

using Sanger sequencing. In the family with ER-positive breast cancer WES identified a rare missense mutation R385 in a DNA mismatch repair gene, RAD50 (rs139372231), in both the index patient and her mother. The RAD50 R385C missense mutation was predicted to be damaging by only one of the three bioinformatics prediction tools used. To help determine the potential clinical relevance of this rare variant in the South African population, an additional 160 breast cancer patients and 301 female controls (101 individuals above the age of 65 years and 200 individuals selected from the genomics database resource) were screened for RAD50 R385C. This resulted in the identification of RAD50 R385C in one relatively young control individual, aged 36 years, with no family history of cancer. In an attempt to clarify the clinical relevance of RAD50 R385C in the index patient with familial breast cancer and her mother, a blood sample was obtained for mutation screening in an affected first-degree cousin as the only remaining affected member in this BRCA-negative family. This patient diagnosed with invasive lobular carcinoma (ILC) at the age of 42 years tested negative for RAD50 R385C. Notably, the index patient was diagnosed with invasive ductal carcinoma (IDC) at the age of 29 years, while her 57-year old mother had IDC in the left breast and invasive lobular carcinoma (ILC) in the right breast.

Further screening for gene variants that may explain clinical heterogeneity within the index family revealed a rare missense mutation Q67E in the Mucin 1 gene (MUC1, chr1: 155160668), predicted to be damaging by all three bioinformatics tools applied. This missense mutation was confirmed by Sanger sequencing in the HER2-positive index case with early-onset breast cancer (<30 years), while absent in her HER2-negative mother with postmenopausal breast cancer and the cousin. Neither the RAD50 nor the MUC1 missense mutations were identified in the exomes of the TNBC patient and the three control individuals. This TNBC patient was furthermore evaluated for potentially causative mutations in the BRCA1 and BRCA2 genes based on tumour pathology, as TNBC is frequently associated with mutations in these high-penetrance genes irrespective of family history. Although a high impact protein truncating mutation was not identified in this patient, two functional BRCA1 polymorphisms (rs16941, E1038G and rs1799966, S1613C) considered potential contributory factors to the disease profile were identified in the TNBC patient. This finding, together with confirmation of at least 10 non-pathogenic BRCA1/2 polymorphic variants previously identified in the index case, confirmed that WES performed in this study on the Proton NGS apparatus generates sufficient coverage of the BRCA1 and BRCA2 genes to replace single gene-by-gene sequencing, provided that any potentially causative mutations identified are verified by Sanger sequencing and family screening where possible.

Failure to identify the rare RAD50 and MUC1 mutations identified in the index case with IDC also in her mother with bilateral breast cancer (only RAD50 identified) and the affected cousin with ILC, may reflect the clinical variability in the family. Although further studies are warranted to screen for causative mutations in the cousin using WES, it is noteworthy that this patient tested positive for one copy of the MTHFR 677 C>T mutation. Since this mutation causing dysfunction of the methylation pathway was detected in both mother and daughter prior to WES, the likelihood that MTHFR 677 C>T contributes to the disease profile in this family cannot be excluded. This multi-functional SNP has previously been linked to ER-positive breast cancer, the only clinical feature shared by all three patients with familial breast cancer, while absent in the TNBC patient. Extended analysis of SNPs in the folate pathway revealed the presence of additional functional variants in the MTR and MTRR genes in the family. Assessment of allelic combinations of MTHFR 677 C>T, MTHFR 1298 A>C, MTR 2756 A>G and MTRR 66 A>G revealed that the TAGG combination is significantly associated with ER-positive breast cancer ($p=0.034$), after adjustment for age, ethnicity and BMI. Wishart (2015) contended that many of the vast number of cancer genes/mutations in humans affect one-carbon metabolism as one of the three major metabolic pathways involved in the pathogenesis of cancer. Differences in detection of the MTRR SNP were also noted when the results obtained with variant calling using hg19 were compared with the MARS, possibly due to differences in allocation of the minor/risk allele according to the MAF as previously determined in world populations. Discrepancies between hg19 and the MARS were also noted for the D148E SNP in the DNA mismatch repair gene, APEX1 (rs1130409). The APEX1 T-allele confirmed in this study to be the minor allele in Caucasians, while it is the major allele in most other world populations studied worldwide, was correctly identified by the MARS but has been missed by hg19. However, the gene effect does not depend on the frequency of a specific allele in a population, as evidenced by the fact that the APEX1 G-allele has clinical relevance in patients receiving radiotherapy.

6.1. Ethical considerations

The advent of new genomic technologies has changed the way researchers approach both basic and translational research. Various questions have arisen due to the vast amounts of information generated by approaches such as exome sequencing, including how to manage, interpret and prioritise this information. Other questions relate to whether or not participating individuals should be allowed access to their sequencing information or whether it should be shared with insurance companies. While the best-defined risk is most insurable, it is impossible to cover risk that cannot be assessed. Appropriate risk classification is essential to maintain a viable voluntary insurance industry since all insurers make decisions based on available genetic

information, including family history and prior diagnostic tests performed in the applicant and his/her family. A common misconception is that genetic testing for breast cancer may result in exclusion from life insurance, however, mounting evidence is accelerating change in this regard and proving testing to be more beneficial (Kotze et al. 2004).

The use of NGS technologies can lead to unexpected findings and complicate clinical utility and patient benefit. This area of genomics still remains a challenge and research is more frequently addressing these questions (Biesecker 2010). The most critical issue regarding commercialization of personalized medicine is the protection of patients. One of the major issues is the fear and potential consequences for patients who are predisposed after genetic testing or found to be non-responsive towards certain treatments. This includes the psychological effects on patients due to genetic testing results. The implications for certain ethnic groups and presence of a common susceptibility allele due to a founder effect also has to be considered (Van der Merwe and van Rensburg 2009; Reeves et al. 2004). As part of the PSGT strategy, the chronic disease screen including a lifestyle intervention plan is purposed to address these issues by providing genetic counseling as a standard of care (Kotze et al. 2005; Kotze et al. 2004).

In the 36-year old breast cancer-free participant of the chronic disease risk screen in whom the RAD50 R385C mutation was identified, the study consent form guided decisions surrounding the management of WES results. The consent form states that the genetic results obtained (e.g. if additional WES was performed) will be made known to the individual only if results indicate that the study participant 1) has a particular disorder or family history relevant to the test results, 2) have a condition or predisposition to developing a condition that is treatable or avoidable e.g. by a lifestyle modification, or 3) may need genetic counselling. This control individual had no family history of cancer and did not receive genetic counselling. This was due to the fact that we could not ascertain the clinical relevance of the RAD50 R385C mutation with the use of various bioinformatics prediction tools that indicated contradictory results.

6.2. Conclusion

A multidisciplinary approach to chronic disease risk management captures the essence of personalized medicine as it could assist clinicians in the development and timely implementation of individual lifestyle-based intervention strategies aimed at decreasing cumulative cardio-metabolic risk in patients with or at risk for breast cancer. In this study we demonstrated the potential value of the CVD multi-gene assay positioned alongside the assessment of relevant non-genetic data such as BMI, to guide clinical and therapeutic decision making across diagnostic boundaries. In addition to its added value in the context of chronic disease risk

screening, the clinical application of our PSGT approach could also assist clinicians in the identification of patients set to derive optimal benefit from extended genetic testing including BRCA screening and ultimately WES in genetically uncharacterised patients or those experiencing drug side effects or failure. A genomic solution is provided to facilitate the lowering of cumulative risk and at the same time identify the need for further analysis beyond the limitations of the CVD multi-gene assay.

The genomic database utilized in this study proved to be a valuable resource for validation of variants detected by WES as well as the interpretation of those of uncertain clinical significance. Its on-going development provides an important resource for health outcomes studies and allows for follow-up and careful monitoring of patients over time. Ultimately, it is envisioned that the patient database which enabled extensive local research to date will continue to inform clinical decision making and optimize delivery of a comprehensive and truly individualized genomics-based healthcare service for breast cancer patients based on disease patterns relevant to the South African population.

WES has the potential to accelerate the discovery of new susceptibility genes which may provide a means to identify novel therapeutic targets to optimize patient care based partly on the assessment of genetic risk factors. Nevertheless, this era of vast data generation and identification of novel genes, while opening new roads in cancer susceptibility, still needs to be treated with caution, as the gaps in knowledge highlighted in this study have indisputable implications for data translation in the clinical setting. Genetic counselling for novel or unexpected findings may present a challenge that needs to be addressed using a multidisciplinary approach which incorporates all relevant information in relation to current and future risk in patients as well as their family members. Inclusion of MARS in the workflow effectively reduced the number of false variants identified by hg19 used as an intermediate bioinformatics step for data translation between the laboratory and clinic. Our findings support previous WES results indicating that the majority of genetically uncharacterised breast cancer may be caused by a combination of low-penetrance mutations exerting their effect against a high-risk backdrop. WES applied beyond the limited scope of single-gene testing has the potential to detect both known and novel mutations across diagnostic boundaries to facilitate prevention of cumulative risk underlying the development of breast cancer and associated comorbidities.

This study led to the development of a EPA for WES that involves 1) IHC to determine ER status and family history of cancer, 2) use of a clinical and lifestyle questionnaire to determine risk factors for CVD and other treatment-related co-morbidities, 3) selection of a gene panel incorporating high-penetrance founder mutations in the BRCA1 and 2 genes as well as low-

intermediate risk alleles implicated in cancer development, recurrence risk and treatment response across the diagnostic spectrum. The genomics database developed as part of this study proved to be a valuable resource for interpretation of variants of uncertain clinical significance. Inclusion of MARS in the workflow effectively reduced the number of false variants when using hg19 as an intermediate bioinformatics step for data translation between the laboratory and clinic. At loci where the major alleles of disease associated variants are absent from hg19, potential causative variants and functional polymorphisms may be missed. Our findings support previous WES results indicating that the majority of genetically uncharacterised familial breast cancer may be caused by a combination of low-moderate penetrance mutations exerting their effect in a high-risk environment such as a high BMI and suboptimal diet. WES enables identification of genetic risk factors of relevance to both cancer development and tailored therapeutic intervention in a single test.

CHAPTER 7

REFERENCES

AbdRaboh NR, Shehata HH, Ahmed MB, Bayoumi FA (2013) HER1 R497K and HER2 I655V polymorphisms are linked to development of breast cancer. Dis Markers 34: 407-417

Abraham JE, Maranian MJ, Driver KE, et al (2010) CYP2D6 gene variants: association with breast cancer specific survival in a cohort of breast cancer patients from the United Kingdom treated with adjuvant tamoxifen. Breast Cancer Res. 12: R64

Abuzeid WM, Jiang X, Shi G, et al (2009) Molecular disruption of RAD50 sensitizes human tumour cells to cisplatin-based chemotherapy. J Clin Investg 119: 1974-1985

Adams PC and Barton JC (2007) Haemochromatosis. Lancet 370: 1855-1860

Ades F, Zardavas D, Bozovic-Spasojevic I, et al (2014) Luminal B Breast Cancer: Molecular Characterization, Clinical Management, and Future Perspectives. JCO 54: 1870.

Adzhubei I, Jordan DM, Sunyaev SR (2013) Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet. 7: 7-20

Adzhubei IA, Schmidt S, Peshkin L, et al (2010) A method and server for predicting damaging missense mutations. Nat Methods 7: 248-249

Agenbag GM (2005) Molecular genetic analysis of familial breast cancer in South Africa. Dissertation, Stellenbosch University. http://hdl.handle.net/10019/953

Agnoli C, Grioni S, Sieri S et al (2015) Metabolic syndrome and breast cancer risk: A case-cohort study nested in a Multicentre Italian Cohort. PLoS ONE. doi:10.1371/journal.pone.0128891

Ahles TA, Saykin AJ, Furstenberg CT, et al (2002) Neuropsychologic impact of standard-dose systemic chemotherapy in long-term survivors of breast cancer and lymphoma. J Clin Oncol 20: 485-493.

Ahmed S, Thomas G, Ghoussaini M, et al (2009) Newly discovered breast cancer susceptibility loci on 3p24 and 17q23. Nature Genetics 41: 585-590

Alberti KG, Zimmet P, Shaw J (2006) MetS - a new world-wide definition. A Consensus Statement from the International Diabetes Federation. Diabet Med. 23: 469-480

Al-Ejeh F, Smart CE, Morriosn BJ, et al (2011) Breast cancer stem cells: Treatment resistance and therapeutic opportunities. Carcinogenesis 32: 650-658

Alexander CM, Landsman PB, Teutsch SM, Haffner SM (2003) NCEP-Defined Metabolic Syndrome, Diabetes, and Prevalence of Coronary Heart Disease Among NHANES III Participants Age 50 Years and Older. Diabetes 52: 1210-1214

Almén MS, Jacobsson JA, Moschonis G et al (2012) Genome wide analysis reveals association of a FTO gene variant with epigenetic changes. Genomics 99: 132-137

Aloraifia F, Bolandb MR, Greenc AJ, Geraghtyb JG (2015) Gene analysis techniques and susceptibility gene discovery in non-BRCA1/BRCA2 familial breast cancer. Surgical Oncology 24: 100-109

American Cancer Society. Cancer Facts & Figures 2015. Atlanta: American Cancer Society, Inc. http://www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2015/

American Society of Clinical Oncology policy statement update (2003) Genetic testing for cancer susceptibility. J Clin Oncol 21: 2397-2406

Ames B (1983) Dietary carcinogens and anticarcinogens. Oxygen radicals and degenerative diseases. Science 221: 1256-1264

Amin DN, Sergina N, Ahuja D, et al (2010) Resiliency and vulnerability in the HER2-HER3 tumourigenic driver. Science Translational Medicine 2: 16ra7. doi: 10.1126/scitranslmed.3000389

Andreassi MG, Botto N, Maffei S (2006) Factor V Leiden, prothrombin G20210A substitution and hormone therapy: indications for molecular screening. Clin Chem Lab Med 44: 514-521

Antoniou AC and Easton DF (2006) Models of genetic susceptibility to breast cancer. Oncogene 25: 5898-5905

Antoniou AC, Spurdle AB, Sinilnikova OM, et al (2008) Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers. Am J Hum Genet. 82: 937-948

Anttila T, Helkala EL, Viitanen M et al (2004) Alcohol drinking in middle age and subsequent risk of mild cognitive impairment and dementia in old age: a prospective population based study. BMJ 329: 539-542

Arendas K, Qiu Q, Gruslin A (2008) Obesity in pregnancy: pre-conceptional to postpartum consequences. J Obstet Gynaecol Can. 30: 477-488

Ariel IM (1987) Breast cancer, a historic review: is the past prologue? In: Ariel IM, Cleary JB, Eds. Breast Cancer Diagnosis and Treatment. New York: McGraw-Hill 3-26

Arneth B, Shams M, Hiemke C, Hartter S (2009) Rapid and reliable genotyping procedure for detection of alleles with mutations, deletion, or/and duplication of the CYP2D6 gene. Clin Biochem 42: 1282-1290

Arpino G, Bardou VJ, Clark GM, Elledge RM (2004). "Infiltrating lobular carcinoma of the breast: tumour characteristics and clinical outcome". Breast Cancer Res. 6: R149-R156

Arun B, Bayraktar S, Liu DD, et al (2011) Response to neoadjuvant systemic therapy for breast cancer in BRCA mutation carriers and noncarriers: a single-institution experience. J Clin Oncol. 29: 3739-3746

Babraham Bioinformatics. http://www.bioinformatics. bbsrc.ac.uk/projects/fastqc. Accessed June 2015

Bamshad MJ, Ng SB, Bigham AW, et al (2011) Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet. 12: 745-755

Bansal V (2010) A statistical method for the detection of variants from next-generation resequencing of DNA pools. Bioinformatics 26: i318-i324

Bapat A, Fishel ML, Kelley MR (2009) Going Ape as an approach to cancer therapeutics. Antioxid Redox Signal 11: 651-667

Bardia A, Arieas ET, Zhang Z et al (2012) Comparison of breast cancer recurrence risk and cardiovascular disease incidence risk among postmenopausal women with breast cancer. Breast Cancer Res Treat 131: 907-914

Bartkova J, Tommiska J, Oplustilova L (2008) Aberrations of the MRE11–RAD50–NBS1 DNA damage sensor complex in human breast cancer: MRE11 as a candidate familial cancer-predisposing gene. Molecular Oncology 2: 296-316

Baselga J and Swain SM (2009) Novel anticancer targets: revisiting HER2 and discovering HER3. Nature Reviews Cancer 9: 463-475

Baselga J, Albanell J, Molina MA, Arribas J (2001) Mechanism of action of trastuzumab and scientific update. Seminars in Oncology 28: 4-11

Bauduer F and Lacombe D (2005) Factor V Leiden, prothrombin 20210A, methylenetetrahydrofolate reductase 677T, and population genetics. Mol Genet Metab. 86: 91-99

Bauer KR, Brown M, Cress RD, et al (2007) Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. Cancer 109: 1721-1728

Beadle G, Baade P, Fritschi L (2009) Acute myeloid leukemia after breast cancer: a population-based comparison with hematological malignancies and other cancers. Ann Oncol *20:*103-109

Beck P, Wysowski DK, Downey W, Butler-Jones D (2003) Statin use and the risk of breast cancer. J Clin Epidemiol 56: 280-285

Beilby J, Ingram D, Hahnel R, Rossi E (2004) Reduced breast cancer risk with increasing serum folate in a case-control study of the C677T genotype of the methylenetetrahydrofolate reductase gene. Eur J Cancer 40*:* 1250-1254

Bender CM, Paraska KK, Sereika SM, et al (2001) Cognitive function and reproductive hormones in adjuvant therapy for breast cancer: a critical review. J Pain Symptom Manage 21: 407-424

Bernard S, Neville KA, Nguyen AT, Flockhart DA (2006) Interethnic differences in genetic polymorphisms of CYP2D6 in the U.S. population: clinical implications. Oncologist 11:126-135

Bernstein C, Prasad AR, Nfonsam V, Bernstein H (2013) DNA Damage, DNA Repair and Cancer, New Research Directions in DNA Repair, Prof. Clark Chen (Ed.), ISBN: 978-953-51-1114-6, InTech, DOI: 10.5772/53919. Available from: http://www.intechopen.com/books/new-research-directions-in-dna-repair/dna-damage-dna-repair-and-cancer

Bertucci F, Finetti P, Cervera N, et al (2008) How basal are triple-negative breast cancers? Int J Cancer 123: 236-240

Berx G, Cleton-Jansen AM, Nollet F, et al (1995) E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. EMBO J 14: 6107-6115

Bianchi F, Raponi M, Piva F, et al (2011) An intronic mutation in MLH1 associated with familial colon and breast cancer. Fam Cancer 10: 27-35

Biesecker LG (2010) Exome sequencing makes medical genomics a reality. Nat. Genet. 42: 13-14

Bird A (2002) DNA methylation patterns and epigenetic memory. Genes Dev. 16: 6-21

Bird BR and Swain SM (2008) Cardiac toxicity in breast cancer survivors: review of potential cardiac problems. Clin Cancer Res 14: 14-24

Bodmer W and Tomlinson I (2010) Rare genetic variants and the risk of cancer. Curr Opin Genet Dev. 20: 262-267

Bogaarts MP, Den Oudsten BL, Roukema JA, et al (2011) The Psychosocial Distress Questionnaire-Breast Cancer (PDQ-BC) is a useful instrument to screen psychosocial problems. Support Care Cancer. 20: 1659-1665

Bogdanova N, Feshchenko S, Schürmann P, et al (2008) Nijmegen Breakage Syndrome mutations and risk of breast cancer. International Journal of Cancer 122: 802-806

Boland JF, Chung CC, Roberson D, et al (2013) The new sequencer on the block: comparison of Life Technology's Proton sequencer to an Illumina HiSeq for whole-exome sequencing. Hum Genet 132: 1153-1163

Bonfrate L, Wang D Q-H, Garruti G, Portincasa P (2014) Obesity and the risk and prognosis of gallstone disease and pancreatitis. Best Practice & Research Clinical Gastroenterology 28: 623-635

Bordeleau L, Panchal S, Goodwin P (2010) Prognosis of BRCA-associated breast cancer: a summary of evidence. Breast Cancer Res Treat 119:13-24

Borges S, Desta Z, Jin Yet al (2010) Composite functional genetic and comedication CYP2D6 activity score in predicting tamoxifen drug exposure among breast cancer patients. J Clin Pharmacol 50: 450-458

Borges S, Desta Z, Li L et al (2006) Quantitative effect of CYP2D6 genotype and inhibitors on tamoxifen metabolism: implication for optimization of breast cancer treatment. Clin Pharmacol Ther 80: 61-74

Bosch N, Junyent N, Gadea N (2012) What factors may influence psychological well-being at three months and one year post BRCA genetic result disclosure? Breast 21: 755-760

Bosch TM, Meijerman I, Beijnen JH, Schellens JH (2006) Genetic polymorphisms of drug-metabolising enzymes and drug transporters in the chemotherapeutic treatment of cancer. Clin Pharmacokinet. 45: 253-285

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32: 314-331

Bouwens CSH, van Rensburg SJ, de Kock L, et al (2012) Influence of genetic factors on the development of breast cancer in the older woman. Current Aging Science 5:140-147

Boyle P, Boniol M, Koechlin A, et al (2012) Diabetes and breast cancer risk: a meta-analysis. Br J Cancer 107: 1608-1617

Bradford LD (2002) CYP2D6 allele frequency in European Caucasians, Asians, Africans and their descendants. Pharmacogenomics 3: 229-243

Brauch H, Schroth W, Goetz MP, et al (2013) Tamoxifen Use in Postmenopausal Breast Cancer: CYP2D6 Matters. Journal of Clinical Oncology 31: 176-180

Breastcancer.org. http://www.breastcancer.org/symptoms/diagnosis/staging. Accessed August 2015

Breasted JH (1930) Ed. The Edwin Smith Surgical papyrus. Chicago, Illinois: The University Chicago Press, special edition: 1984

Brezden CB, Phillips KA, Abdolell M, et al (2000) Cognitive function in breast cancer patients receiving adjuvant chemotherapy. J Clin Oncol 18: 2695-2701

Broeks A, Urbanus JH, Floore AN, et al (2000) ATM-heterozygous germline mutations contribute to breast cancer-susceptibility. Am J Hum Genet. 66: 494-500

Brugarolas J, Lei K, Hurley RL, et al (2004) Regulation of mTOR function in response to hypoxia by REDD1 and the TSC1/TSC2 tumour suppressor complex. Genes Dev. 18: 2893-2904

Bryant HE, Petermann E, Schultz N, et al (2009) PARP is activated at stalled forks to mediate Mre11- dependent replication restart and recombination. EMBO J. 28: 2601-2615

Bugianesi E, Leone N, Vanni E, et al (2002) Expanding the natural history of nonalcoholic steatohepatitis: from cryptogenic cirrhosis to hepatocellular carcinoma. Gastroenterology 123: 134-140

Byrski T, Gronwald J, Huzarski T, et al (2009) Neoadjuvant therapy with cisplatin in BRCA1-positive breast cancer patients. J Clin Oncol. 9: A4

Cadoo KA, Fornier MN, Morris PG (2013) Biological subtypes of breast cancer: current concepts and implications for recurrence patterns. Quarterly Journal of Nuclear Medicine and Molecular Imaging 57: 312-321

Cai Q, Shu X, Wen W, et al (2004) Genetic polymorphism in the manganese superoxide dismutase gene, antioxidant intake, and breast cancer risk: results from the Shanghai Breast Cancer Study. Breast Cancer Res 6: R647-R655

Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ (2003) Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N Engl J Med. 348: 1625-1638

Campbell CD, Chong JX, Malig M, et al (2012) Estimating the human mutation rate using autozygosity in a founder population. Nature Genetics 44: 1277-1281

Campeau PM, Foulkes WD, Tischkowitz MD (2008) Hereditary breast cancer: New genetic developments, new therapeutic avenues. Human Genetics 124: 31-42

Cantley LC (2002) The phosphoinositide 3-kinase pathway. Science 296: 1655-1657

Capasso I, Esposito E, de Laurentiis NM, et al (2014) Metabolic syndrome-breast cancer link varies by intrinsic molecular subtype. Diabetology & Metabolic Syndrome 6: 105

Carey LA, Perou CM, Livasy CA, et al (2006) Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA 295: 2492-2502

Cauley JA, Lucas FL, Kuller LH, et al (1999) Elevated serum estradiol and testosterone concentrations are associated with a high risk for breast cancer. Study of Osteoporotic Fractures Research Group. Ann Intern Med. 130: 270-277

Cauley JA, McTiernan A, Rodabough RJ, et al. (2006) Statin use and breast cancer: Prospective results from the Women's Health Initiative. J Natl Cancer Inst 98: 700-707

Cella D, Fallowfield L, Barker P, et al (2006) Quality of life of postmenopausal women in the ATAC ("Arimidex", tamoxifen, alone or in combination) trial after completion of 5 years' adjuvant treatment for early breast cancer. Breast Cancer Res Treat. 100: 273-284

Chambon P, Ramuz M, Doly J (1964) Relation between soluble DNA- dependent RNA polymerase and “aggregate” RNA polymerase. Biochem. Bio- phys. Res. Commun. 21: 156-161

Chan M, Ji SM, Yeo ZX, et al (2012) Development of a next-generation sequencing method for BRCA mutation screening: a comparison between a high-throughput and a benchtop platform. J Mol Diagn 14: 602-612

Chandler WM, Rowe LR, Florell SR, et al (2012) Differentiation of Malignant Melanoma From Benign Nevus Using a Novel Genomic Microarray With Low Specimen Requirements. Arch Pathol Lab Med. 136: 947-955

Chang S, Pollack LM, Colditz GA (2013) Obesity, Mortality, and Life Years Lost Associated With Breast Cancer in Nonsmoking US Women, National Health Interview Survey, 1997–2000. Prev Chronic Dis 10: 130112

Chen J, Gammon MD, Chan W, et al (2005) One-carbon metabolism, MTHFR polymorphisms, and risk of breast cancer. Cancer Res 65: 1606-1614

Chen S, Parmigiani G (2007) Meta-analysis of BRCA1 and BRCA2 penetrance. J Clin Oncol 25:1329-1333

Cheon JY, Mozersky J, Cook-Deegan R (2014) Variants of uncertain significance in BRCA: a harbinger of ethical and policy issues to come? Genome Medicine 6: 121

Chikarmane SA, Tirumani SH, Howard SA, et al (2015) Metastatic patterns of breast cancer subtypes: What radiologists should know in the era of personalized cancer medicine. Clinical Radiology 70: 1-10

Chiusolo P, Reddiconto G, Casorelli I, et al (2002) Preponderance of methylenetetrahydrofolate reductase C677T homozygosity among leukemia patients intolerant to methotrexate. Ann Oncol 13:1915-1918

Chlebowski RT (2011) obesity adversely affects survival in postmenopausal patients. JCO 30: 126-128

Choi M, Scholl UI, Ji W, et al (2009) Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci USA. 106: 19096-19101

Choi SW and Mason JB (2002) Folate status: effects on pathways of colorectal carcinogenesis. J Nutr 132: 2413S-2418S

Choi Y and Chan AP (2015) PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics. doi: 10.1093/bioinformatics/btv195

Choi Y, Sims GE, Murphy S, et al (2012) Predicting the Functional Effect of Amino Acid Substitutions and Indels. PLoS ONE 7: e46688

Chua W, Goldstein D, Lee CK, et al (2009) Molecular markers of response and toxicity to FOLFOX chemotherapy in metastatic colorectal cancer. Br. J. Cancer 101: 998-1004

Chun S and Fay JC (2009) Identification of deleterious mutations within three human genomes. Genome Res. 19: 1553-1561

Church GM (2006) Genomes for all. Sci. Am. 294: 46-54

Citri A, Skaria KB, Yarden Y (2003) The deaf and the dumb: the biology of ErbB-2 and ErbB-3. Experimental Cell Research 284: 54-65

Claus EB, Risch N, Thompson WD (1991) Genetic analysis of breast cancer in the cancer and steroid hormone study. The American Journal of Human Genetics 48: 232-242

Cleator S, Ashworth A (2004) Molecular profiling of breast cancer: clinical implications. Br J Cancer 90: 1120-1124

Coate L, Cuffe S, Horgan A, et al (2010) Germline genetic variation, cancer outcome, and pharmacogenetics. J Clin Oncol 28: 4029-4037

Cohen HJ, Lan L, Archer L, Kornblith AB (2012) Impact of age, comorbidity and symptoms on physical function in long-term breast cancer survivors. J Geriatr Oncol. 3: 82-89

Cohen V, Panet-Raymond V, Sabbaghian N, et al (2003) Methylenetetrahydrofolate reductase polymorphism in advanced colorectal cancer: a novel genomic predictor of clinical response to fluoropyrimidine-based chemotherapy. Clin Cancer Res 9:1611-1615

Collaborative Group on Hormonal Factors in Breast Cancer (2002). Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50 302 women with breast cancer and 96973 women without the disease. Lancet 360: 187-195

Coppen A, Bolander-Gouaille C (2005) Treatment of depression: time to consider folic acid and vitamin B12. J Psychopharmacol 19:59-65

Correa Geyer F and Reis-Filho JS (2009) Microarray-based gene expression profiling as a clinical tool for breast cancer management: are we there yet? Int J Surg Pathol 17: 285-302

Cox A, Dunning AM, Garcia-Closas M, et al (2007) A common coding variant in CASP8 is associated with breast cancer risk. Nature Genetics 39: 352-358

Cronin-Fenton DP and Lash TL (2011) Clinical epidemiology and pharmacology of CYP2D6 inhibition related to breast cancer outcomes. Expert Rev Clin Pharmacol. 4: 363-377

Curtis C, Shah SP, Chin S, et al (2012) The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 486: 346-352

Cushman M, Kuller LH, Prentice R, et al (2004) Estrogen Plus Progestin and Risk of Venous Thrombosis. JAMA. 292: 1573-1580

Dai X1, Fagerholm R, Khan S, et al (2015) INPP4B and RAD50 have an interactive effect on survival after breast cancer. Breast Cancer Res Treat. 149: 363-371

Damiola F, Pertesi M, Oliver J, et al (2014) Rare key functional domain missense substitutions in MRE11A, RAD50, and NBN contribute to breast cancer susceptibility: results from a Breast Cancer Family Registry case-control mutation-screening study. Breast Cancer Research 16: R58.

Dancey J (2010) mTOR signaling and drug development in cancer. Nat Rev Clin Oncol. 7: 209-219

Davis and Kaklamani VG (2012) MetS and Triple-Negative Breast Cancer: A New Paradigm. International Journal of Breast Cancer. doi:10.1155/2012/809291

Davis W, van Rensburg SJ, Cronje FJ (2013) The fat mass and obesity-associated FTO rs9939609 polymorphism is associated with elevated homocysteine levels in patients with multiple sclerosis screened for vascular risk factors. Metab Brain Dis. 29: 409-419

de Magalhães JP, Finch CE, Janssens G (2010) Next-generation sequencing in aging research: emerging applications, problems, pitfalls and possible solutions. Ageing Research Reviews 9: 315-323

De Mattia E and Toffoli G (2009) C677T and A1298C MTHFR polymorphisms, a challenge for antifolate and fluoropyrimidine-based therapy personalisation. Eur J Cancer. 45:1333-1351

de Medina P, Payré BL, Bernad J et al (2004) tamoxifen is a potent inhibitor of cholesterol esterification and prevents the formation of foam cells. J Pharmacol Exp Ther. 308: 1165-1173

De Moulin D (1983) A short history of breast cancer. Boston: Martinus Nijhoff 1-107

de Villiers JNP, Hillerman R, De Jong G, et al (1999) High prevalence of the CYS282TYR mutation facilitates an improved diagnostic service for haemochromatosis in South Africa. SAMJ 89: 279-282

de Vogel S, Wouters KA, Gottschalk RW, et al (2009) Genetic variants of methyl metabolizing enzymes and epigenetic regulators: associations with promoter CpG island hypermethylation in colorectal cancer. Cancer Epidemiol. Biomarkers Prev. 18: 3086-3096

de Wit LM, van Straten A, van Herten M, Penninx BW, Cuijpers P (2009) Depression and body mass index, a u-shaped association. BMC Public Health 9: 14

Decensi A, Maisonneuve P, Rotmensz N, et al (2005) Effect of Tamoxifen on Venous Thromboembolic Events in a Breast Cancer Prevention Trial. Circulation 111: 650-656

Degner JF, Marioni JC, Pai AA, Pickrell JK, NkadoriE, et al (2009) Effect of read-mapping biases on detecting allele-specific expression from RNA sequencing data. Bioinformatics 25: 3207-3212

Del Re M, Michelucci A, Simi P, Danesi R. (2012) Pharmacogenetics of anti-estrogen treatment of breast cancer. Cancer Treat Rev. 38: 442-450

Delport D, Schoeman R, van der Merwe N, et al (2014) Significance of dietary folate intake, homocysteine levels and MTHFR 677 C>T genotyping in South African patients diagnosed with depression: test development for clinical application. Metab Brain Dis 29: 377-384

D'Empaire I, Guico-Pabia CJ, Preskorn SH (2011) Antidepressant treatment and altered CYP2D6 activity: are pharmacokinetic variations clinically relevant? J Psychiatr Pract. 17: 330-339

den Dunnen JT and Antonarakis SE (2000) Mutation Nomenclature Extensions and Suggestions to Describe Complex Mutations: A Discussion. Hum. Mutat. 15: 7-12

Den Heijer M, Lewington S, Clarke R (2005) Homocysteine, MTHFR and risk of venous thrombosis: a meta-analysis of published epidemiological studies. J Thromb Haemost. 3: 292-299

Derwinger K, Wettergren Y, Odin E, et al (2009) A study of the MTHFR gene polymorphism C677T in colorectal cancer. Clin. Colorectal Cancer. 8: 43-48

Deshpande RA, Williams GJ, Limbo O (2014) ATP-driven Rad50 conformations regulate DNA tethering, end resection, and ATM checkpoint signaling. The Emnbo Journal. Doi: 10.1002/embj.201386100

Després JP and Lemieux I (2006) Abdominal obesity and metabolic syndrome. Nature 444: 881-887

Dever SM, White ER, Hartman MC, Valerie K (2012) BRCA1-directed, enhanced and aberrant homologous recombination: Mechanism and potential treatment strategies. Cell Cycle 11: 687-694

Devlin AM, Ngai YF, Ronsley R, Panagiotopoulos C (2012) Cardiometabolic risk and the MTHFR C677T variant in children treated with second-generation antipsychotics. Transl Psychiatry. 2: e71

Dewey FE, Chen R, Cordero SP, et al (2011) Phased whole-genome genetic risk in a family quartet using a major allele reference sequence. PLoS Genet. 7:e1002280

Dewey FE, Grove ME, Pan C, et al (2014) Clinical interpretation and implications of whole-genome sequencing. JAMA 311: 1035-1044

Dey S (2014) Preventing breast cancer in LMICs via screening and/or early detection: The real and the surreal. World J Clin Oncol. 5: 509-519

Di Renzo L, Marsella LT, Sarlo F, et al (2014) C677T gene polymorphism of MTHFR and MetS: response to dietary intervention. J Transl Med 12: 329

Dieci MV, Orvieto E, Dominici M, et al (2014) Rare breast cancer subtypes: histological, molecular, and clinical peculiarities. Oncologist 19: 805-813

Dignam JJ, Wieand K, Johnson KA, et al (2003) Obesity, tamoxifen use, and outcomes in women with estrogen receptor-positive early-stage breast cancer. J Natl Cancer Inst 95: 1467-1476

Dossus L and Benusiglio PR (2015) Lobular breast cancer: incidence and genetic and non-genetic risk factors. Breast Cancer Res. 17: 37

Dowsett M, King N, Dowsett M, et al (1995) In vivo measurement of aromatase inhibition by letrozole (CGS 20267) in postmenopausal patients with breast cancer. Clin Cancer Res. 1: 1511-1515

Dowsett M, Stein RC, Mehta A, Coombes RC (1990) Potency and selectivity of the non-steroidal aromatase inhibitor CGS 16949A in postmenopausal breast cancer patients. Clin Endocrinol (Oxf) 32: 623-634

Drugs.com. http://www.drugs.com/health-guide/breast-cancer.html. Accessed October 2015.

Druley TE, Vallania FLM, Wegner DJ, et al (2009) Quantification of rare allelic variants from pooled genomic DNA. Nat Methods 6: 263-265

Durbin RM, Abecasis GR, Altshuler DL, Auton A, Brooks LD, et al. (2010) A map of human genome variation from population-scale sequencing. Nature 467: 1061-1073

Duval ST and Weedie R (2000) Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 562: 455-463

Early Breast Cancer Trialists' Collaborative Group (EBCTCG) (2005) Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 365: 1687-1717

Easton DF, Pharoah PDP, Antoniou AC (2015) Gene-Panel Sequencing and the Prediction of Breast-Cancer Risk. N Engl J Med 372: 2243-2257

Easton DF, Pooley KA, Dunning AM, et al (2007) Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 447: 1087-1093

Eberling JL, Wu C, Tong-Turnbeaugh R, Jagust WJ (2000) Estrogen- and tamoxifen-associated effects on brain structure and function. Neuroimage 21: 364-371

Eheman CR, Shaw KM, Ryerson AB, et al (2009). The changing incidence of in situ and invasive ductal and lobular breast carcinomas: United States, 1999-2004. Cancer Epidemiol. Biomarkers Prev. 18: 1763-1769

Ehrlich M (2009) DNA hypomethylation in cancer cells. Epigenomics 1: 239-259

Ehrnhoefer DE, Wong BK, Hayden MR (2011) Convergent pathogenic pathways in Alzheimer's and Huntington's diseases: shared targets for drug development. Nat Rev Drug Discov 10: 853-67

Elhassan HOM and Abdalla MHA (2015) Methylenetetrahydrofolate Reductase (MTHFR C677T) Polymorphism in Sudanese Patients with Deep Vein Thrombosis. Int. J Biomed Res 6: 323-326

Ellis MJ, Ding L, Shen D, et al (2012) Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature 486:353-360

Elvers I, Johansson F, Groth P, et al (2011). UV stalled replication forks restart by re-priming in human fibroblasts. Nucleic Acids Res. 39: 7049-7057

Emmert-Streib F, de Matos Simoes R, Mullan P (2014) The gene regulatory network for breast cancer: integrated regulatory landscape of cancer hallmarks. Frontiers in Genetics. doi: 10.3389/fgene.2014.00015

Endogenous Hormones and Breast Cancer Collaborative Group (2002) Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. J Natl Cancer Inst. 94: 606-616

Ericson U, Sonestedt E, Gullberg Bo, et al (2007) High folate intake is associated with lower breast cancer incidence in postmenopausal women in the Malmö Diet and Cancer cohort. Am J Clin Nutr 86: 434-443

Eroles P, Bosch A, Pérez-Fidalgo JA, Lluch A (2012) Molecular biology in breast cancer: Intrinsic subtypes and signaling pathways. Cancer Treatment Reviews 38: 698-707

Esteva FJ and Hortobagyi GN (2006) Comparative assessment of lipid effects of endocrine therapy for breast cancer: implications for cardiovascular disease prevention in postmenopausal women. Breast 15: 301-312

Etienne MC, Formento JL, Chazal M, et al (2004) Methylenetetrahydrofolate reductase gene polymorphisms and response to fluorouracil-based treatment in advanced colorectal cancer patients. Pharmacogenetics. 14: 785-792

Evans A (2004) Ductal carcinoma in situ (DCIS): are we overdetecting it? Breast Cancer Research 6: P23

Evans WE and Relling MV (2004) Moving towards individualized medicine with pharmacogenomics. Nature 429: 464-468

Ewald B, Sampath D, Plunkett W (2008) ATM and the Mre11-Rad50-Nbs1 complex respond to nucleoside analogue-induced stalled replication forks and contribute to drug resistance. Cancer Res 68: 7947-7955

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-Calling of Automated Sequencer Traces Using Phred. I. Accuracy Assessment. Genome Res. 8: 175-185

Faith MS, Matz PE, Jorge MA (2002) Obesity-depression associations in the population. J Psychosom Res 534: 935-942

Falleti MG, Sanfilippo A, Maruff P, et al (2005) The nature and severity of cognitive impairment associated with adjuvant chemotherapy in women with breast cancer: a meta-analysis of the current literature. Brain Cogn 59: 60-70

Feigelson HS, Jonas CR, Robertson AS, et al (2003) Alcohol, folate, methionine, and risk of incident Breast cancer in the American Cancer Society Cancer Prevention Study II Nutrition Cohort. Cancer Epidemiol Biomarkers Prev 12: 161-164

Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, et al (2013) Cancer incidence and mortality patterns in Europe: Estimates for 40 countries in 2012. European Journal of Cancer 49: 1374-1403

Feron VJ, Til HP, de Vrijer F, et al (1991) Aldehydes: occurrence, carcinogenic potential, mechanism of action and risk assessment. Mutat Res 259: 363-385

Ferraldeschi R and Newman WG (2010) The Impact of CYP2D6 Genotyping on Tamoxifen Treatment. Pharmaceuticals 3: 1122-1138

Figueiredo JC, Brooks JD, Conti DV, et al (2011) Risk of contralateral breast cancer associated with common variants in BRCA1 and BRCA2: potential modifying effect of BRCA1/BRCA2 mutation carrier status. Breast Cancer Res Treat. 127: 819-829

Fisher B, Bauer M, Margolese R, et al (1985) Five-year results of a randomized clinical trial comparing total mastectomy and segmental mastectomy with or without radiation in the treatment of breast cancer. N Engl J Med 312: 665-673

Fisher B, Costantino JP, Wickerham DL, et al (1998) Tamoxifen for Prevention of Breast Cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. J Natl Cancer Inst. 90: 1371-1388

Fisher B, Redmond C, Poisson R, et al (1989) Eight-year results of a randomized clinical trial comparing total mastectomy and lumpectomy with or without irradiation in the treatment of breast cancer. N Engl J Med 320: 822-828

Fisher B, Costantino JP, Wickerham DL, et al (2005) Tamoxifen for the prevention of breast cancer: current status of the National Surgical Adjuvant Breast and Bowel Project P-1 study. J Natl Cancer Inst. 97: 1652-1662

Fisher LRF (2011) Evaluation of high-throughput methodology for multi-gene screening in patients with non-alcoholic fatty liver disease (NAFLD). http://scholar.sun.ac.za/handle/10019.1/17896

Flanagan SE, Patch AM, Ellard S (2010) Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. Genet Test Mol Biomarkers 14: 533-537

Flegal KM, Graubard BI, Williamson DF, Gail MH (2005) Excess deaths associated with underweight, overweight, and obesity. JAMA 293: 1861-1867

Ford NA, Rossi EL2, Barnett K, et al (2015) Omega-3-Acid Ethyl Esters Block the Protumourigenic Effects of Obesity in Mouse Models of Postmenopausal Basal-like and Claudin-Low Breast Cancer. Cancer Prev Res. 8: 796-806

Fostira F, Tsitlaidou M, Papadimitriou C, et al (2012) Prevalence of BRCA1 mutations among 403 women with triple-negative breast cancer: implications for genetic screening selection criteria: a Hellenic Cooperative Oncology Group Study. Breast Cancer Res Treat. 134: 353-362

Foulkes WD, Stefansson IM, Chappuis PO, et al (2003) Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. J Natl Cancer Inst. 95: 1482-1485

Franklin MC, Carey KD, Vajdos FF, et al (2004) Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. Cancer Cell 5: 317-328

Frazer DM and Anderson GJ (2003) The orchestration of body iron intake: how and where do enterocytes receive their cues? Blood Cells Mol Dis. 30: 288-297

Fredholm H, Eaker S, Frisell J, et al (2009) Breast cancer in young women: poor survival despite intensive treatment. PLoS One 4: e7695

Freedman RA and Partridge AH (2013) Management of breast cancer in very young women. Breast. 22: S176-S179

Friso S, Choi SW, Girelli D, et al (2002) A common mutation in the 5,10-methylenetetrahydrofolate reductase gene affects genomic DNA methylation through an interaction with folate status. Proc Natl Acad Sci USA 99:5606-5611

Frosst P, Blom HJ, Milos R, et al (1995) A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. Nat Genet. 10: 111-113

Frudakis T, Thomas M, Ginjupalli S et al (2007) CYP2D6*4 polymorphism is associated with statin-induced muscle effects. Pharmacogenet Genomics 17: 695-707

Frueh FW, Amur S, Mummaneni P, et al (2008) Pharmacogenomic biomarker information in drug labels approved by the United States food and drug administration: prevalence of related drug use. Pharmacotherapy 28: 992-998

Fulford LG, Easton DF, Reis-Filho JS, et al (2006) Specific morphological features predictive for the basal phenotype in grade 3 invasive ductal carcinoma of breast. Histopathology 49: 22-34

Furgason JM and Bahassi el M (2013) Targeting DNA repair mechanisms in cancer. Pharmacol. Ther. 137: 298-308

Futreal PA, Coin L, Marshall M, et al (2004) A census of human cancer genes. Nature reviews Cancer 4: 177-183

Gaedigk A, Coetsee C (2008) The CYP2D6 gene locus in South African coloureds: unique allele distributions, novel alleles and gene arrangements. Eur J Clin Pharmacol 64: 465-475

Gahl WA, Markello TC, Toro C, et al (2012) The National Institutes of Health Undiagnosed Diseases Program : insights into rare diseases. Genet Med. 14: 51-59

Gaiddon C, Moorthy NC, Prives C (1999) Ref-1 regulates the transactivation and pro-apoptotic functions of p53 in vivo. EMBO J 18: 5609-5621

Gann PH and Morrow M (2003) Combined Hormone Therapy and Breast Cancer: A Single-Edged Sword. JAMA. 289: 3304-3306

Garber JE, Halabi S, Tolaney SM et al (2010) Factor V Leiden Mutation and Thromboembolism Risk in Women Receiving Adjuvant Tamoxifen for Breast Cancer. J Natl Cancer Inst. 102: 942-949

García JM, Silva J, Peña C, et al (2004) Promoter methylation of the PTEN gene is a common molecular change in breast cancer. Genes Chromosomes and Cancer 41: 117-124

Garcia M, Jemal A, Ward EM, et al (2007) Global Cancer Facts & Figures 2007. Atlanta, GA: American Cancer Society

Geisler J, King N, Anker G, et al (1996) Influence of anastrozole (Arimidex), a selective, non-steroidal aromatase inhibitor, on in vivo aromatisation and plasma oestrogen levels in postmenopausal women with breast cancer. Br J Cancer 74: 1286-1291

Geisler J, King N, Anker G, et al (1998) In vivo inhibition of aromatization by exemestane, a novel irreversible aromatase inhibitor, in postmenopausal breast cancer patients. Clin Cancer Res. 4: 2089-2093

Genetics home reference. http://ghr.nlm.nih.gov/gene/RAD50. Accessed 7 May 2015

Genin E, Hannequin D, Wallon D, et al (2011) APOE and Alzheimer disease: a major gene with semi-dominant inheritance. Mol Psychiatry 16: 903-907

Ghayad SE and Cohen PA (2010) Inhibitors of the PI3K/Akt/mTOR pathway: new hope for breast cancer patients. Recent Pat Anticancer Drug Discov. 5: 29-57

Ghosh R, Narasanna A, Wang SE, et al (2011) Trastuzumab has preferential activity against breast cancers driven by HER2 homodimers. Cancer Research 71: 1871-1882

Giam M and Rancati G (2015) Aneuploidy and chromosomal instability in cancer: a jackpot to chaos. Cell Division 10: 3

Giordano SH, Cohen DS, Buzdar AU, et al (2004). Breast carcinoma in men: a population-based study. Cancer 101: 51-57

Glück S, de Snoo F, Peeters J, et al (2013) Molecular subtyping of early-stage breast cancer identifies a group of patients who do not benefit from neoadjuvant chemotherapy. Breast Cancer Res Treat. 139: 759-767

Goetz MP, Knox SK, Suman VJ et al (2007) The impact of cytochrome P450 metabolism in women receiving adjuvant tamoxifen. Breast Cancer Res Treat. 101: 113-121

Goetz MP, Rae JM, Suman VJ, et al (2005) Pharmacogenetics of tamoxifen biotransformation is associated with clinical outcomes of efficacy and hot flashes. J. Clin. Oncol. 23: 9312-9318

Gonzalez-Angulo AM, Morales-Vasquez F, Hortobagyi GN (2007) Overview of resistance to systemic therapy in patients with breast cancer. Adv Exp Med Biol. 608: 1-22

Goodman JE, Lavigne JA, Wu K, et al (2001) COMT genotype, micronutrients in the folate metabolic pathway and breast cancer risk. Carcinogenesis 22:1661-1665

Goyette P, Sumner JS, Milos R, et al (1994) Human methylenetetrahydrofolate reductase: isolation of cDNA, mapping and mutation identification. Nat Genet. 7: 195-200

Gracia-Aznarez FJ, Fernandez V, Pita G, et al (2013) Whole exome sequencing suggests much of non-BRCA1/BRCA2 familial breast cancer is due to moderate and low penetrance susceptibility alleles. PLoS One 8: e55681

Grant KA, Apffelstaed JP, Wright C, et al (2013) MammaPrint Pre-screen Algorithm (MPA) reduces chemotherapy in patients with early-stage breast cancer. South African Medical Journal 103: 522-526

Grant KA, Pienaar FM, Brundyn K, et al (2015) Incorporating microarray assessment of HER2 status in clinical practice supports individualised therapy in early-stage breast cancer. Breast 24: 137-142

Graus-Porta D, Beerli RR, Daly JM, Hynes NE (1997) ErbB-2, the preferred heterodimerization partner of all ErbB receptors, is a mediator of lateral signalling. EMBO Journal 16: 1647-1655

Graziano F, Kawakami K, Ruzzo A et al (2006) Methylenetetrahydrofolate reductase 677C/T gene polymorphism, gastric cancer susceptibility and genomic DNA hypomethylation in an at-risk Italian population. Int J Cancer 118: 628-632

Grilli S (2006) Tamoxifen (TAM): the dispute goes on. Ann Ist Super Sanita 42: 170-173

Grotle M, Hagen KB, Natvig B, et al (2008) Obesity and osteoarthritis in knee, hip and/or hand: an epidemiological study in the general population with 10 years follow-up. BMC Musculoskelet Disord. 9: 132

Gruvberger S, Ringner M, Chen Y, et al (2001) Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. Cancer Res 61: 5979-5984

Guillotin D and Martin SA (2014) Exploiting DNA mismatch repair deficiency as a therapeutic strategy. Exp. Cell Res. 329: 110-115

Gurrin LC, Osborne NJ, Constantine CC, et al (2008) The natural history of serum iron indices for HFE C282Y homozygosity associated with hereditary hemochromatosis. Gastroenterology 135: 1945-1952

Hack TF, Pickles T, Ruether JD, et al (2010) Predictors of distress and quality of life in patients undergoing cancer therapy: impact of treatment type and decisional role. Psychooncology 19: 606-616

Hall JM, Lee MK, Newman B, et al (1990) Linkage of early-onset familial breast cancer to chromosome 17q21. Science 250: 1684-1689

Hall N (2007) Advanced sequencing technologies and their wider impact in microbiology. J. Exp. Biol. 209: 1518-1525

Halsted CH, Villanueva JA, Devlin AM, Chandler CJ (2002) Metabolic interactions of alcohol and folate. J Nutr 132: 2367S-2372S

Hamann U (2000) Hereditary breast cancer: high risk genes, genetic testing and clinical implications. Clin Lab. 46: 447-461

Hammoud AO, Gibson M, Peterson CM, et al (2008) Impact of male obesity on infertility: a critical review of the current literature. Fertil. Steril. 90: 897-904

Hara K, Yonezawa K, Weng QP, et al (1998) Amino acid sufficiency and mTOR regulate p70 S6 kinase and eIF-4E BP1 through a common effector mechanism. J Biol Chem. 273: 14484-14494

Harmon DL, Shields DC, Woodside JV, et al (1999) Methionine synthase D919G polymorphism is a significant but modest determinant of circulating homocysteine concentrations. Genet. Epidemiol. 17: 298-309

Harrison SE, Watson EK, Ward AM, et al (2011) Primary health and supportive care needs of long-term cancer survivors: a questionnaire survey. J Clin Oncol 29: 2091-2098

Hartge P, Struewing JP, Wacholder S, et al (1999) The prevalence of common BRCA1 and BRCA2 mutations among Ashkenazi Jews. Am. J. Hum. Genet. 64: 963-970

Harvie M, Howell A, Vierkant RA et al (2005) Association of gain and loss of weight before and after menopause with risk of postmenopausal breast cancer in the Iowa women's health study. Cancer Epidemiol Biomarkers Prev. 14: 656-661

Hassold JT, Burrage LC, Chan ER, et al (2001) Maternal folate polymorphisms and the etiology of human nondisjunction, Am J Hum Genet. 9: 434-43

Healey CS, Dunning AM, Teare MD, et al (2000) A common variant in BRCA2 is associated with both breast cancer risk and prenatal viability. Nat Genet. 26: 362-364

Hegde ML, Izumi T, Mitra S (2012) Oxidized base damage and single-strand break repair in mammalian genomes: role of disordered regions and posttranslational modifications in early enzymes. Prog Mol Biol Transl Sci. 110: 123-153

Heijmans BT, Boer JM, Suchiman HE, et al (2003) A common variant of the methylenetetrahydrofolate reductase gene (1p36) is associated with an increased risk of cancer. Cancer Res. 63:1249-1253

Heikkinen K, Karppinen SM, Soini Y, et al (2003) Mutation screening of Mre11 complex genes: indication of RAD50 involvement in breast and ovarian cancer susceptibility. J Med Genet. 40: e131

Heikkinen K, Rapakko K, Karppinen SM, et al (2006) RAD50 and NBS1 are breast cancer susceptibility genes associated with genomic instability. Carcinogenesis 27: 1593-1599

Hemminki A, Avizienyt E, Roth S, et al (1998) A serine/threonine kinase gene defective in Peutz-Jeghers syndrome. Nature 114: 667-668

Hertz DL, McLeod HL, Irvin WJ Jr (2012) Tamoxifen and CYP2D6: A Contradiction of Data. Oncologist. Oncologist. 17: 620-630

Hickey M, Elliott J, Davison SL (2012) Hormone replacement therapy. BMJ 344: e763

Hindorff LA, MacArthur J, Morales J, et al (2013) European Bioinformatics Institute. A Catalog of Published Genome-Wide Association Studies. Available at: www.genome.gov/gwastudies.

Hira B, Pegoraro RJ, Rom L, Moodley J (2003) Absence of Factor V Leiden, thrombomodulin and prothrombin gene variants in Black South African women with pre-eclampsia and eclampsia. BJOG 110: 327-328

Hobbs CA, Sherman SL, Yi P, et al (2000) Polymorhpisms in genes involved in folate metabolism as maternal risk factors for Down syndrome. Am J Hum Genet. 67: 623-630

Hollestelle A, Wasielewski M, Martens JW, Schutte M (2010) Discovering moderate-risk breast cancer susceptibility genes. Curr Opin Genet Dev. 20: 268-276

Holmström P, Marmur J, Eggertsen G et al (2002) Mild iron overload in patients carrying the HFE S65C gene mutation: a retrospective study in patients with suspected iron overload and healthy controls. Gut. 51: 723-730

Holzman D (2009) Tamoxifen, antidepressants, and CYP2D6: the conundrum continues. J Natl Cancer Inst 101:1370-1371

Homer N, Merriman B, Nelson SF (2009) Bfast: an alignment tool for large scale genome resequencing. PLoS One 4: e7767

Homer. Iliad. (1966) Translated by WHD Rouse. New York: A Signet Classic. New American Library, 36

Hopper JL, Jenkins MA, Dowty JG, et al (2012) Using tumour pathology to identify people at high genetic risk of breast and colorectal cancers. Pathology 44: 89-98

Hortobagyi GN (1988) Treatment of breast cancer. N Engl J Med 339: 974-984

Horwich A, Dearnaley DP, Norman A, et al. (1994) Accelerated chemotherapy for poor prognosis germ cell tumours. Eur J Cancer 30A: 1607-1611

Hsieh P and Yamane K (2008) DNA mismatch repair: molecular mechanism, cancer, and ageing. Mech. Ageing Dev. 129: 391-407

Hu J, Zhou GW, Wang N, Wang YJ (2010) MTRR A66G polymorphism and breast cancer risk: a meta-analysis. Breast Cancer Res Treat. 124: 779-784

Hu JJ, Smith TR, Miller MS, et al (2001) Amino acid substitution variants of APE1 and XRCC1 genes associated with ionizing radiation sensitivity. Carcinogenesis 22: 917-922

Huang ZP, Hankinson SE, Colditz GA, et al (1997) Dual effects of weight and weight gain on breast cancer risk. JAMA 278: 1407–1411

Hubner RA, Muir KR, Liu JF, et al (2006) Folate metabolism polymorphisms influence risk of colorectal adenoma recurrence.Cancer Epidemiol. Biomarkers Prev. 15: 1607-1613

Hughes JB, Berger C, Rodland MS, et al (2009) Pertuzumab increases epidermal growth factor receptor down-regulation by counteracting epidermal growth factor receptor-ErbB2 heterodimerization. Molecular Cancer Therapeutics 8: 1885-1892

Hunter DJ, Kraft P, Jacobs KB, et al (2007) A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nature Genetics 39: 870-874

Hurtado A, Holmes KA, Geistlinger TR, et al (2008) Regulation of ERBB2 by oestrogen receptor-PAX2 determines response to tamoxifen. Nature 456: 663-666

Iglesias-Linares A, Yañez-Vico RM, González-Moles MA (2010) Potential role of HDAC inhibitors in cancer therapy: insights into oral squamous cell carcinoma. Oral Oncol. 46: 323-329

International Agency for Research on Cancer (2013) Latest world cancer statistics. global cancer burden rises to 14.1 million new cases in 2012: marked increase in breast cancers must be addressed. In: WHO International Agency for Research on Cancer, ed. Lyon, France and Geneva, Switzerland: International Agency for Research on Cancer and World Health Organization, 1

Jacques PF, Bostom AG, Williams RR, et al (1996) Relation between folate status, a common mutation in methylenetetrahydrofolate reductase, and plasma homocysteine concentrations. Circulation 93: 7-9

Jakubowska A, Gronwald J, Menkiszak J, et al (2007) Methylenetetrahydrofolate reductase polymorphisms modify BRCA1-associated breast and ovarian cancer risks. Breast Cancer Res Treat 104:299-308

Jansen MPHM, Foekens JS, van Staveren IL, et al (2005) Molecular Classification of Tamoxifen-Resistant Breast Carcinomas by Gene Expression Profiling. J Clin Oncol 23: 732-740

Jenkins EO, Deal AM, Anders CK, et al (2014) Age-Specific Changes in Intrinsic Breast Cancer Subtypes: A Focus on Older Women. Oncologist 19: 1076-1083

Jerusalem G, Rorive A, Collignon J (2014) Use of mTOR inhibitors in the treatment of breast cancer: an evaluation of factors that influence patient outcomes. Breast Cancer (Dove Med Press) 6: 43-57

Jim HS, Phillips KM, Chait S et al (2012) Meta-analysis of cognitive functioning in breast cancer survivors previously treated with standard-dose chemotherapy. J Clin Oncol. 30: 3578-3587

Jim HS, Small BJ, Minton S, et al (2012) History of major depressive disorder prospectively predicts worse quality of life in women with breast cancer. Ann Behav Med 43: 402-408

Jin Y, Desta Z, Stearns V, et al (2005) CYP2D6 genotype, antidepressant use, and tamoxifen metabolism during adjuvant breast cancer treatment. J. Natl. Cancer Inst. 97: 30-39

Jiricny J (2006) The multifaceted mismatch-repair system. Nat. Rev. Mol. Cell Biol. 7: 335-346

Jordan VC, Collins MM, Rowsby L, Prestwich G (1977) A monohydroxylated metabolite of tamoxifen with potent antioestrogenic activity. J Endocrinol 75: 305-316

Kabat GC, Kim M, Phipps AI (2011) Smoking and alcohol consumption in relation to risk of triple-negative breast cancer in a cohort of postmenopausal women. Cancer Causes Control 22: 775-783

Kabuto M, Akiba S, Stevens RG, et al (2000) A prospective study of estradiol and breast cancer in Japanese women. Cancer Epidemiol Biomarkers Prev. 9: 575-579

Kahvejian A, Quackenbush J, Thompson JF (2008) What would you do if you could sequence everything? Nature Biotechnology 26: 1125-1133

Kainu T, Juo SHH, Desper R, et al (2000) Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. PNAS 97: 9603-9608

Kakkoura MG, Demetriou CA, Loizido MA, et al (2015) Single-nucleotide polymorphisms in one-carbon metabolism genes, Mediterranean diet and breast cancer risk: a case-control study in the Greek-Cypriot female population. Genes Nutr 10: 453

Kaklamani V, Yi N, Sadim M, et al (2011) The role of the fat mass and obesity associated gene (FTO) in breast cancer risk. BMC Med Genet 12: 52-62

Kallianpur AR, Hall LD, Yadav M (2004) Increased Prevalence of the HFE C282Y Hemochromatosis Allele in Women with Breast Cancer. Cancer Epidemiol Biomarkers Prev 13: 205

Kamangar F, Dores GM, Anderson WF (2006) Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. J Clin Oncol. 24: 2137-2150

Kang S, Kim JW, Kang GH, et al (2005) Polymorphism in folate- and methionine-metabolizing enzyme and aberrant CpG island hypermethylation in uterine cervical cancer. Gynecol. Oncol. 96: 173-180

Karp SE, Tonin PN, Bégin LR, et al (1997) Influence of BRCA1 mutations on nuclear grade and estrogen receptor status of breast carcinoma in Ashkenazi Jewish women. Cancer 80: 435-441

Kaufmann M, von minckwitz G, Bear HD, et al (2006) Recommendations from an international expert panel on the use of neoadjuvant (primary) systemic treatment of operable breast cancer: new perspectives. Annals of Oncology 18: 1927-1934

Keijzer M, den Heijer BM, Blom HJ, et al (2002) Interaction between hyperhomocysteinemia, mutated methylenetetrahydrofolatereductase (MTHFR) and inherited thrombophilic factors in recurrent venous thrombosis. Thromb Haemost. 88: 723-728

Kelland L (2007) Broadening the clinical use of platinum drug–based chemotherapy with new analogues. Expert Opin. Investig. Drugs 16: 1009-1021

Khatcheressian JL, Wolff AC, Smith TJ, et al (2006) American Society of Clinical Oncology 2006 Update of the Breast Cancer Follow-Up and Management Guidelines in the Adjuvant Setting. Journal of Clinical Oncology 24: 5091-5097

Khoury MJ, Gwinn M, Ioannidis JP (2010) The emergence of translational epidemiology: from scientific discovery to population health impact. Am J Epidemiol. 172: 517-524

Khoury MJ, Gwinn M, Yoon PW et al (2007) The continuum of translation research in genomic medicine: how can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? Genet Med. 9:665-674

Kim S, Lee M, Lee K et al (2010) Use of antidepressants in patients with breast cancer taking tamoxifen. J Breast Cancer 13:325-336

Kim YI (1999) Folate and carcinogenesis: evidence, mechanisms, and implications. J Nutr Biochem. 10:66-88

Kim YI (2006) Folate: a magic bullet or a double edged sword for colorectal cancer prevention? Gut 55: 1387-1389

King MC, Marks JH, Mandell JB, the New York Breast Cancer Study Group (2003) Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. Science 302: 643-646

Kinoshita E, van Rossum-Fikkert S, Sanchez H, et al (2015) Human RAD50 makes a functional DNA-binding complex. Biochimie 113: 47-53

Kiyotani K, Mushiroda T, Imamura CK et al (2010) Significant effect of polymorphisms in CYP2D6 and ABCC2 on clinical outcomes of adjuvant tamoxifen therapy for breast cancer patients. J Clin Oncol 28: 1287-1293

Klerk M, Lievers KJ, Kluijtmans LA, et al (2003) The 2756A>G variant in the gene encoding methionine synthase: its relation with plasma homocysteine levelsand risk of coronary heart disease in a Dutch case-control study. Thromb Res. 110: 87-91

Klerk M, Verhoef P, Clarke R, et al (2002) MTHFR 677C>T polymorphism and risk of coronary heart disease: a meta-analysis. J Am Med Assoc 288:2023-2031

Koboldt DC, Chen K, Wylie T, et al (2009) VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics 25: 2283-2285

Kofuji S, Kimura H, Nakanishi H, et al (2015) INPP4B Is a PtdIns(3,4,5)P3 Phosphatase that can act as a tumor suppressor. Cancer discovery 5: 730-739

Kopelman P (2007) Health Risks Associated with Overweight and Obesity. Obesity Reviews 8:13-17

Kotsopoulos J, Kim YI, Narod SA (2012) Folate and breast cancer: what about high-risk women? Cancer Causes Control 23:1405-1420

Kotze MJ and Thiart R (2003) Genetics of dyslipidaemia. CME J 21: 399-402

Kotze MJ and van Rensburg SJ (2012) Pathology supported genetic testing and treatment of cardiovascular disease in middle age for prevention of Alzheimer's disease. Metab Brain Dis. 27: 255-266

Kotze MJ van Rensburg SJ (2012) Pathology supported genetic testing and treatment of cardiovascular disease in middle age for prevention of Alzheimer's disease. Metab Brain Dis 27: 255-266

Kotze MJ, Kriegshäuser G, Thiart R, et al (2003) Simultaneous detection of multiple familial hypercholesterolaemia mutations facilitates an improved diagnostic service in South African patients at high risk of cardiovascular disease. Mol Diagn 7: 169-174

Kotze MJ, Lückhoff HK, Brand T, et al (2015) Apolipoprotein E ε-4 as a genetic determinant of Alzheimer's disease heterogeneity. Degenerative Neurological and Neuromuscular Disease 5: 9-18

Kotze MJ, Lückhoff HK, Peeters AV, et al (2015) Genomic medicine and risk prediction across the disease spectrum. Crit Rev Clin Lab Sci 19: 1-18

Kotze MJ, Malan J, Pienaar R, Apffelstaedt J (2005) The role of molecular genetic testing in modern breast health management. S Afr Fam Pract. 47: 38-40

Kotze MJ, Schorn D, Coetzer P (2004) The impact of genetic testing on life insurance. J Genomics Afr Soc. 1: 1-11

Kotze MJ, van Velden DP, Botha K, et al (2013) Pathology-supported genetic testing directed at shared disease pathways for optimized health in later life. Personalized Med. 10: 497-507

Koushik A, Kraft P, Fuchs CS, et al (2006) Nonsynonymous polymorphisms in genes in the one-carbon metabolism pathway and associations with colorectal cancer. Cancer Epidemiol. Biomarkers Prev. 15: 2408-2417

Krauss RM, Winston M, Fletcher RN, Grundy SM (1998) Obesity: impact of cardiovascular disease. Circulation 98: 1472-1476

Krejsa C, Rogge M, Sadee W (2006) Protein therapeutics: new applications for pharmacogenetics. Nature Reviews Drug Discovery 5: 507-521

Kriege M, Brekelmans CT, Boetes C, et al (2004) Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition. New England Journal of Medicine 351: 427-437

Krontiris TG, Devlin B, Karp DD, et al (1993) An association between the risk of cancer and mutations in the HRAS1 minisatellite locus. N. Engl. J. Med. 329: 517-523

Kufe DW (2013) MUC1-C oncoprotein as a target in breast cancer: activation of signaling pathways and therapeutic approaches. Oncogene 32: 1073-1081

Kuhl CK, Schmutzler RK, Leutner CC, et al (2000) Breast MR imaging screening in 192 women proved or suspected to be carriers of a breast cancer susceptibility gene: preliminary results. Radiology 215: 267-279

Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature Protocols 4: 1073-1081

Kwan ML, Kroenke CH, Sweeney C, et al (2015) Association of high obesity with PAM50 breast cancer intrinsic subtypes and gene expression. BMC Cancer 15: 278

Lajous M, Romieu I, Sabia S et al (2006) F. Folate, vitamin B(12) and postmenopausal Breast cancer in a prospective study of French women. Cancer Causes Control 17: 1209-1213

Lakhani SR, Reis-Filho JS, Fulford L, et al (2005) Prediction of BRCA1 status in patients with breast cancer using estrogen receptor and basal phenotype. Clin Cancer Res 11: 5175-5180

Lalloo F and Evans DG (2012) Familial breast cancer. Clin Genet. 82: 105-11

Lambrinoudaki I, Papadimitriou D, Kaparos G, et al (2013) MTHFR C677T polymorphism modifies the effect of HRT on metabolic parameters in postmenopausal women. Climacteric 16: 568-575

Lander ES and Botstein D (1986) Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphisms. Proc Natl Acad Sci USA 83: 7353-7357

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10: R25

Lash TL, Pedersen L, Cronin-Fenton D (2008) Tamoxifen's protection against breast cancer recurrence is not reduced by concurrent use of the SSRI citalopram. Br J Cancer 99: 616-621

Laumann EO, Kang JH, Glasser DB, et al (2008) Lower urinary tract symptoms are associated with depressive symptoms in white, black and Hispanic men in the United States. J Urol. 180: 233-240

Leclerc D, Campeau E, Goyette P, et al (1996) Human methionine synthase: cDNA cloning and identification of mutations in patients of the cblG complementation group of folate/cobalamin disorders. Hum. Mol. Genet. 5: 1867-1874

Lee-Hoeflich ST, Crocker L, Yao E, et al (2008) A central role for HER3 in HER2-amplified breast cancer: implications for targeted therapy. Cancer Research 68: 5878-5887

Lengacher CA, Reich RR, Paterson CL et al (2015) The effects of mindfulness-based stress reduction on objective and subjective sleep parameters in women with breast cancer: a randomized controlled trial. Psycho-Oncology 24: 424-432

Leong ASY and Zhuang Z (2011) The changing role of pathology in breast cancer diagnosis and treatment. Pathobiology 78: 99-114

Lesko LJ and Zineh I (2010) DNA, drugs and chariots: on a decade of pharmacogenomics at the US FDA. Pharmacogenomics 11: 507-512

Levy-Lahad E, Catane R, Eisenberg S, et al (1997) Founder BRCA1 and BRCA2 mutations in Ashkenazi Jews in Israel: frequency and differential penetrance in ovarian cancer and in breast-ovarian cancer families. Am J Hum Genet. 60: 1059-1067

Lewison EF (1953) The surgical treatment of breast cancer; an historical and collective review. Surgery 34: 904-953

Li CI, Daling JR, Porter PL et al (2009) Relationship between potentially modifiable lifestyle factors and risk of second primary contralateral breast cancer among women diagnosed with estrogen receptor-positive invasive breast cancer. J Clin Oncol. 27: 5312-5318

Li H and Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754-1760

Li H, Handsaker B, Wysoker A, et al (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078-2079

Li J, Chen Y, Wu H, Li L (2014) Apolipoprotein E (Apo E) gene polymorphisms and recurrent pregnancy loss: a meta-analysis. J Assist Reprod Genet 31: 139-148

Li LC, Carroll PR, Dahiya R (2005) Epigenetic changes in prostate cancer: implication for diagnosis and treatment. J. Natl. Cancer Inst. 97: 103-115

Li M, Lu L-Y, Yang C-Y, Wang S, Yu X (2013) The FHA and BRCT domains recognize ADP-ribosylation during DNA damage response. Genes Dev. 27: 1752-1768

Li Y and Agarwal P (2009) A pathway-based view of human diseases and disease relationships. PLoS One 4: e4346

Li Y, Yu WH, Ren J, et al (2003c) Heregulin targets γ-catenin to the nucleolus by a mechanism dependent on the DF3/MUC1 protein. Mol Cancer Res. 1: 765-775

Liang S, He L, Zhao X, et al (2011) MicroRNA Let-7f Inhibits Tumour Invasion and Metastasis by Targeting MYH9 in Human Gastric Cancer. PLoS ONE 6: e18409

Liberopoulos E, Karabina SA, Tselepis A, et al (2002) Are the effects of tamoxifen on the serum lipid profile modified by apolipoprotein E phenotypes? Oncology 62: 115-120

Lindström LS, Karlsson E, Wilking UM, et al (2012) Clinically used breast cancer markers such as estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2 are unstable throughout tumour progression. J Clin Oncol. 30: 2601-2608

Lippman SM and Brown PH (1999) Tamoxifen Prevention of Breast Cancer: an Instance of the Fingerpost. JNCI J Natl Cancer Inst. 91: 1809-1819

Liu G, Zhu H, Lagou V, et al (2010) FTO variant rs9939609 is associated with body mass index and waist circumference, but not with energy intake or physical activity in European- and African American youth. BMC Med Genet. 11: 57

Liu X, Jian X, Boerwinkle E (2011) dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. Hum Mutat. 32: 894-899

Livasy CA, Karaca G, Nanda R, et al (2006) Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. Mod Pathol 19: 264-271

Lohmann PL, Rao ML, Ludwig M, et al (2001) Influence of CYP2D6 genotype and medication on the sparteine metabolic ratio of psychiatric patients. Eur J Clin Pharmacol. 57: 289-295

Lomen N, Johannsson O, Bendahl PO, et al (1998) Steroid receptors in hereditary breast carcinomas associated with BRCA1 or BRCA2 mutations or unknown susceptibility genes. Cancer 83*:* 310-319

Longnecker MP (1994) Alcoholic beverage consumption in relation to risk of breast cancer: meta-analysis and review. Cancer Causes Contr. 5:73-82

Lorizio W, Rugo H, Beattie MS et al (2011) Pharmacogenetic testing affects choice of therapy among women considering tamoxifen treatment. Genome Med. 3: 64-75

Louie SM, Roberts LS, Nomura DK (2013) Mechanisms linking obesity and cancer. Biochim Biophys Acta. 1831: 1499-1508

Luckhoff HK, Brand T, van Velden DP (2015) Clinical relevance of apolipoprotein E genotyping based on a family history of Alzheimer's disease. Curr Alzheimer Res. 12: 210-217

Lund MJ, Trivers KF, Porter PL, et al (2009) Race and triple negative threats to breast cancer survival: a population-based study in Atlanta, GA. Breast Can Res Treat. 113: 357-370

Luppino FS, de Wit LM, Bouvy PF, et al (2010) Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. Arch Gen Psychiatry 67: 220-229

Lupski JR, Reid JG, Gonzaga-Jauregui C, et al (2010) Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. N Engl J Med. 362: 1181-1191

Lyon GJ, Wang K (2012) Identifying disease mutations in genomic medicine settings: current challenges and how to accelerate progress. Genome Med. 4: 1-16

Ma AD and Udden MM (2007) Iron metabolism, iron overload, and the porphyrias. Kahn MJ, Gregory SA. American Society of Hematology Self-Assessment Program. Washington, DC: American Society of Hematology 61-77

Maae E, Andersen RF, Steffensen KD, et al (2012) Prognostic impact of VEGFA germline polymorphisms in patients with HER2-positive primary breast cancer. Anticancer Res. 32: 3619-3627

Macis D, Maisonneuve P, Johansson H, et al (2007) Methylenetetrahydrofolate reductase (MTHFR) and breast cancer risk: a nested-case-control study and a pooled meta-analysis. Breast Cancer Res Treat 106: 263-271

Mackay A, Weigett B, Grigoriadis A (2011) Microarray based class discovery for molecular classification of breast cancer: Analysis of interobserver agreement. J Natl Cancer Inst. 103: 662-673

MacNeill FA, Jones AL, Jacobs S, et al (1992) The influence of aminoglutethimide and its analogue rogletimide on peripheral aromatisation in breast cancer. Br J Cancer 66: 692-697

Magi A, D'Aurizio R, Palombo F, et al (2015) Characterization and identification of hidden rare variants in the human genome. BMC Genomics 16: 340

Malhotra GK, Zhao K, Band H, Band V (2010) Histological, molecular and functional subtypes of breast cancers. Cancer Biol Ther. 10: 955-960

Mann A, Miksys SL, Gaedigk A, et al (2011) The neuroprotective enzyme CYP2D6 increases in the brain with age and is lower in Parkinson's disease patients. Neurobiol Aging 33: 2160-2171

Manolio T, Collins FS, Cox NJ, et al (2009) Finding the missing heritability of complex diseases. Nature 461: 747-753

Manson JE, Hsia J, Johnson KC, et al (2003) Estrogen plus Progestin and the Risk of Coronary Heart Disease. N Engl J Med. 349: 523-534

Mardis ER (2011) A decade's perspective on DNA sequencing technology Nature 470: 198-203

Marjoribanks J, Farquhar C, Roberts H, et al (2012) Long term hormone therapy for perimenopausal and postmenopausal women. Cochrane Database Syst Rev. 7: CD004143

Maruti SS, Ulrich CM, Jupe ER, White E (2009) MTHFR C677T and postmenopausal breast cancer risk by intakes of one-carbon metabolism nutrients: a nested case-control study. Breast Cancer Research 11: R91

Mato JM and Lu SC (2005) Homocysteine, the bad thiol. Hepatology 41: 976-979

Maynard S, Schurman SH, Harboe C, et al (2009) Base excision repair of oxidative DNA damage and association with cancer and aging. Carcinogenesis 30: 2-10

McCabe N, Turner NC, Lord CJ, et al (2006) Deficiency in the repair of DNA damage by homologous recombination and sensitivity to poly (ADP-ribose) polymerase inhibition. Cancer Res. 66: 8109-8115

McCarthy MI and Hirschhorn JN (2008) Genome-wide association studies: potential next steps on a genetic journey. Human Molecular Genetics 17: R156-R165

Mcgee DL (2004) Body mass index and mortality: a meta-analysis based on person-level data from twenty-six observational studies. Ann Epidemiol. 15: 87-97

McKenna A, Hanna M, Banks E, et al (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20: 1297-1303

Medema RH and Macůrek L (2012) Checkpoint control and cancer. Oncogene 31: 2601-2613

Meijers-Heijboer H, van den Ouweland A, Klijn J, et al (2002) Low-penetrance susceptibility to breast cancer due to CHEK2*1100delC in noncarriers of BRCA1 or BRCA2 mutations. Nat Genet. 31: 55-59

Mellick GD (2006) CYP450, genetics and Parkinson's disease: gene x environment interactions hold the key. J Neural Transm Suppl 70: 159-165

Mierla D, Szmal C, Neagos D, et al (2012) Association of Prothrombin (A20210G) and Factor V Leiden (A506G) with Recurrent Pregnancy Loss. Maedica (Buchar) 7: 222-226

Miki Y, Swensen J, Shattuck-Eidens D, et al (1994) A strong candidate for the 17-linked breast and ovarian cancer susceptibility gene BRCA1. Science 266: 66-71

Miller KD and Sledge GW (1999) The role of chemotherapy for metastatic breast cancer. Hematol Oncol Clin North Am. 13: 415-434

Milne RL, Gaudet MM, Spurdle AB, et al (2010) Assessing interactions between the associations of common genetic susceptibility variants, reproductive history and body mass index with breast cancer risk in the breast cancer association consortium: a combined case-control study. Breast Cancer Research 12: R110

Mitrunen K, Kataja V, Eskelinen M, et al (2002) Combined COMT and GST genotypes and hormone replacement therapy associated breast cancer risk. Pharmacogenetics 12: 67-72

Mittal R, Chaudhry N, Pathania S, Mukherjee TK (2014) Mechanistic insight of drug resistance with special focus on iron in estrogen receptor positive breast cancer. Curr Pharm Biotechnol. 15: 1141-1157

Molina MA, Codony-Servat J, Albanell J, et al (2001) Trastuzumab (Herceptin), a humanized antiHER2 receptor monoclonal antibody, inhibits basal and activated HER2 ectodomain cleavage in breast cancer cells. Cancer Research 61: 4744-4749

Moore RJ, Chamberlain RM, Khuri FR (2004) Apolipoprotein E and the risk of breast cancer in African-American and non-Hispanic white women. A review. Oncology 66: 79-93

Moriyama Y, Okamura T, Kajinami K, et al (2002) Effects of serum B vitamins on elevated plasma homocysteine levels associated with the mutation of methylenetetrahydrofolate reductase gene in Japanese. Atherosclerosis 164: 321-328

Morris GJ, Naidu S, Topham AK, et al (2007) Differences in breast carcinoma characteristics in newly diagnosed African-American and Caucasian patients: a single-institution compilation compared with the national cancer institute's surveillance, epidemiology, and end results database. Cancer 110: 876-884

Morton ME (1996) Logarithm of odds (lods) for linkage in complex inheritance. Proc Natl Acad Sci USA 93: 3471-3476

Morton NE (1955) Sequential tests for the detection of linkage. Am J Hum Genet. 7: 277-318

Mountzios G, Aivazi D, Kostopoulos I, et al (2014) Differential expression of the insulin-like growth factor receptor among early breast cancer subtypes. PLoS One 9: e91407

Moysich KB, Freudenheim JL, Baker JA, et al (2000) Apolipoprotein E genetic polymorphism, serum lipoproteins, and breast cancer risk. Mol Carcinog. 27: 2-9

Muendlein A, Hubalek M, Geller-Rhomberg S, et al (2014) Significant survival impact of MACC1 polymorphisms in HER2 positive breast cancer patients. EJC 50: 2134-2141

Muendlein A, Lang AH, Geller-Rhomberg S, et al (2013) Association of a common genetic variant of the IGF-1 gene with event-free survival in patients with HER2-positive breast cancer. J Cancer Res Clin Oncol. 139: 491-498

Nagato LC, de Souza Pinhel MA, de Godoy JM, Souza DR (2012) Association of ApoE genetic polymorphisms with proximal deep venous thrombosis. J Thromb Thrombolysis 33: 116-119

Nagele P and Liggett SB (2011) Genetic variation, β-blockers, and perioperative myocardial infarction. Anesthesiology 115: 1316-1327

Nakamura T, Imai Y, Matsumoto T, et al (2007) Estrogen prevents bone loss via estrogen receptor alpha and induction of Fas ligand in osteoclasts. Cell 130: 811-823

Nakamura Y, Ratain MJ, Cox NJ, et al (2012) Re: CYP2D6 genotype and tamoxifen response in postmenopausal women with endocrine-responsive breast cancer: the Breast International Group 1-98 trial. J Natl Cancer Inst. 104: 1264; author reply 1266-1268

Naushad SM, Pavani A, Rupasree Y, et al (2012) Association of aberrations in one-carbon metabolism with molecular phenotype and grade of breast cancer. Mol Carcinog. 51: E32-E41

Nelen MR, Padberg GW, Peeters EA, et al (1996) Localization of the gene for Cowden disease to chromosome 10q22-23. Nat Genet. 13: 114-116

Nelson HD, Vesco KK, Haney E, et al (2006) Nonhormonal therapies for menopausal hot flashes: systematic review and meta-analysis. JAMA 295: 2057-2071

Newman WG, Hadfield KD, Latif A, et al (2008) Impaired tamoxifen metabolism reduces survival in familial breast cancer patients. Clin Cancer Res. 14: 5913-5918

Ng PC and Henikoff S (2003) SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. 31: 3812-3814

Ng SB, Bigham AW, Buckingham KJ, et al (2010) Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. Nat Genet. 42: 790-793

Ng SB, Bigham AW, Buckingham KJ, et al (2010) Exome sequencing identifies the cause of a mendelian disorder. Nat Genet. 42: 30-35

Ngoma TA (2006) World Health Organization cancer priorities in developing countries. Annals of Oncology 17: viii9-viii14

Nicoloso MS, Sun H, Spizzo R, et al (2010) Single-nucleotide polymorphisms inside microRNA target sites influence tumor susceptibility. Cancer Res. 70: 2789-2798

Noll DM, Mason TM, Miller PS (2006) Formation and Repair of Interstrand Cross-Links in DNA. Chem Rev. 106: 277-301

Norton N, Robertson PD, Rieder MJ, et al (2012) Evaluating pathogenicity of rare variants from dilated cardiomyopathy in the exome era. Circ Cardiovasc Genet. 5: 167-174

Nowell SA, Ahn J, Rae JM, et al (2005) Association of genetic variation in tamoxifen-metabolizing enzymes with overall survival and recurrence of disease in breast cancer patients. Breast Cancer Res. Treat. 91: 249-258

Nseir S, Di Pompeo C, Soubrier S, et al (2005) Effect of ventilator-associated tracheobronchitis on outcome in patients without chronic respiratory failure: a case-control study. Critical Care 9: R238-R245

O'Brien KM, Cole SR, Tse C, et al (2010) Intrinsic breast tumour subtypes, race, and long-term survival in the Carolina Breast Cancer Study. Clin Cancer Res. 16: 6100-6110

O'Rawe J, Jiang T, Sun G, et al (2013) Low concordance of multiple variant-calling pipelines: practical implications for exome and genome sequencing. Genome Medicine 5: 28

Obdeijn IM, Loo CE, Rijnsburger AJ, et al (2010) Assessment of false-negative cases of breast MR imaging in women with a familial or genetic predisposition. Breast Cancer Research and Treatment 119: 399-407

Oliveira KC, Bianco B, Verreschi IT, et al (2008) Prevalence of the polymorphism MTHFR A1298C and not MTHFR C677T is related to chromosomal aneuploidy in Brazilian Turner Syndrome patients, Arq Bras Endocrinol Metabol 52: 1374-1381

Oliver GR (2012) Considerations for clinical read alignment and mutational profiling using next-generation sequencing. Version 2. F1000Res. 1: 2

Olteanu H, Munson T, Banerjee R (2002) Differences in the efficiency of reductive activation of methionine synthase and exogenous electron acceptors between the common polymorphic variants of human methionine synthase reductase. Biochemistry. 41: 13378-13385

Onitilo AA, Engel JM, Greenlee RT, Mukesh BN (2009) Breast cancer subtypes based on ER/PR and Her2 expression: comparison of clinicopathologic features and survival. Clinical Medicine & Research 7: 4-13

O'Rawe J, Jiang T, Sun G, et al (2013) Low concordance of multiple variant-calling pipelines: practical implications for exome and genome sequencing. Genome Medicine 5: 28

Osborne NJ, Gurrin LC, Allen KJ, et al (2010) HFE C282Y homozygotes are at increased risk of breast and colorectal cancer. Hepatology 51: 1311-131

O'Shaughnessy J, Osborne C, Pippen J, et al (2009) Efficacy of BSI-201, a poly (ADP-ribose) polymerase-1 (PARP1) inhibitor, in combination with gemcitabine/carboplatin (G/C) in patients with metastatic triple-negative breast cancer (TNBC): Results of a randomized phase II trial. J Clin Oncol. 27: 18S3

Oyama K, Kawakami K, Maeda K, et al (2004) The association between methylenetetrahydrofolate reductase polymorphism and promoter methylation in proximal colon cancer. Anticancer Res. 24: 649-54

Paganini-Hill A and Clark LJ (2000) Preliminary assessment of cognitive function in breast cancer patients treated with tamoxifen. Breast Cancer Res Treat. 64: 165-76

Pagon RA (2002) Genetic testing for disease susceptibilities: consequences for genetic counselling. Trends Mol Med. 8: 306-307

Papakostas GI, Shelton RC, Zajecka JM, et al (2012) L-methylfolate as adjunctive therapy for SSRI-resistant major depression: results of two randomized, double-blind, parallel-sequential trials. Am J Psychiatry 169: 1267-1274

Paradiso A and Formenti S (2011) Hereditary breast cancer: clinical features and risk reduction strategies. Ann Oncol 22: 131-136

Park D, et al (2014) Rare mutations in RINT1 predispose carriers to breast and Lynch Syndrome-spectrum cancers. Cancer Discovery, doi: 10.1158/2159-8290

Park GY, Wilson JJ, Song Y, Lippard SJ (2012) Phenanthriplatin, a monofunctional DNA-binding platinum anticancer drug candidate with unusual potency and cellular activity profile. Proc. Natl. Acad. Sci. USA 109: 11987-11992

Park JY, Yoo HW, Kim BR, et al (2008) Identification of a novel human Rad51 variant that promotes DNA strand exchange. Nucleic Acids Res. 36: 3226-3234

Park SK, Yoo KY, Lee SJ, et al (2000) Alcohol consumption, glutathione S-transferase M1 and T1 genetic polymorphisms and breast cancer risk. Pharmacogenetics 10: 301-309

Park YH, Kim ST, Cho EY, et al (2010) A risk stratification by hormonal receptors (ER, PgR) and HER-2 status in small (< or = 1 cm) invasive breast cancer: who might be possible candidates for adjuvant treatment? Breast Cancer Res. Treat. 119: 653-661

Parkin DM, Boyd L, Walker LC (2011) The fraction of cancer attributable to lifestyle and envirnmental factors in the UK in 2010. Br J Cancer 05: S77-S78

Partridge AH, Gelber S, Peppercorn J, et al (2008) Fertility and menopausal outcomes in young breast cancer survivors. Clin Breast Cancer 8: 65-69

Partridge AH, Goldhirsch A, Gelber S, et al (2010) Chapter 92: Breast Cancer in Younger Women, in Harris JR, Lippman ME, Morrow M, Osborne CK. Diseases of the Breast, 4th edition, Lippincott Williams & Wilkins

Patel C, Sivadas A, Tabassum R, et al. (2013) Whole genome sequencing in support of wellness and health maintenance. Genome Medicine 5: 58

Patten SB and Barbui C (2004) Drug-induced depression: a systematic review to inform clinical practice. Psychother Psychosom. 73: 207-215

Pattnaik S, Vaidyanathan S, Pooja DG, et al (2012) Customisation of the Exome Data Analysis Pipeline Using a Combinatorial Approach. PLoS ONE 7: e30080

Paz MF, Avila S, Fraga MF, et al (2002) Germ-line variants in methyl-group metabolism genes and susceptibility to DNA methylation in normal tissues and human primary tumours. Cancer Res. 62: 4519-4524

Peerbooms OL, van Os J, Drukker M, et al (2011) Meta-analysis of MTHFR gene variants in schizophrenia, bipolar disorder and unipolar depressive disorder: evidence for a common genetic vulnerability? Brain Behav Immun 25: 1530-1543

Pepe C, Guidugli L, Sensi E, et al (2007) Methyl group metabolism gene polymorphisms as modifiers of breast cancer risk in Italian BRCA1/2 carriers. Breast Cancer Res Treat 103: 29-36

Pepe G, Rickards O, Vanegas OC, et al (1997) Prevalence of factor V Leiden mutation in non-European populations. Thromb Haemost. 77: 329-331

Perez EA, Romond EH, Suman VJ, et al (2011) Four-year follow-up of trastuzumab plus adjuvant chemotherapy for operable human epidermal growth factor receptor 2-positive breast cancer: joint analysis of data from NCCTG N9831 and NSABP B-31. J Clin Oncol. 29: 3366-3373

Perou CM (2011) Molecular Stratification of Triple-Negative Breast Cancers. The Oncologist 16: 61-70

Perou CM, Sorlie T, Eisen MB, et al (2000) Molecular portraits of human breast tumours. Nature 406: 747-752

Petekkaya I, Sahin U, Gezgen G, et al (2013) Association of breast cancer subtypes and body mass index. Tumouri 99: 129-133

Peyrin-Biroulet L, Rodriguez-Guéant RM, Chamaillard M, et al (2007) Vascular and cellular stress in inflammatory bowel disease: revisiting the role of homocysteine. Am. J. Gastroenterol. 102: 1108-1115

Pharmgkb. https://www.pharmgkb.org/pathway/PA145011119). Accessed August 2015

Pharoah PD, Antoniou AC, Easton DF, Ponder BA (2008) Polygenes, risk prediction, and targeted prevention of breast cancer. N Engl J Med. 358: 2796-2803

Phipps AI, Buist DS, Malone KE, et al (2011) Family history of breast cancer in first-degree relatives and triple-negative breast cancer risk. Breast Cancer Res Treat 126: 671-678

Phipps AI, Buist DS, Malone KE, et al (2012) Breast density, body mass index, and risk of tumour marker-defined subtypes of breast cancer. Ann Epidemiol 22: 340-348

Phipps AI, Ichikawa L, Bowles EJA, et al (2010) Defining Menopausal Status in Epidemiologic Studies: A Comparison of Multiple Approaches and their Effects on Breast Cancer Rates. Maturitas. 67: 60-66

Piccart-Gebhart MJ, Procter M, Leyland-Jones B, et al (2005) Herceptin Adjuvant (HERA) Trial Study Team. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. N Engl J Med. 353: 1659-1672

Pike MC, Spicer DV, Dahmoush L, Press MF (1998) Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk. Epidemiol Rev. 15: 17-35

Pinto AC, de Azambuja E (2011) Improving quality of life after breast cancer: dealing with symptoms. Maturitas 70: 343-348

Pollard KS, Hubisz MJ, Rosenbloom KR, et al (2010) Detection of nonneutral substitution rates on mammalian phylogenies Genome Res. 20: 110-121

Porto G and De Sousa M (2007) Iron overload and immunity. World J Gastroenterol. 13: 4707-4715

Prentice RL, Caan B, Chlebowski RT, et al (2006) Low-fat dietary pattern and risk of invasive breast cancer: the Women's Health Initiative Randomized Controlled Dietary Modification Trial. JAMA 295: 629-642

Preskorn SH, Greenblatt DJ, Flockhart D, et al (2007) Comparison of duloxetine, escitalopram, and sertraline effects on cytochrome P450 2D6 function in healthy volunteers. J Clin Psychopharmacol. 27: 28-34

Prospective Studies Collaboration (2009) Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. Lancet 373: 1083-1096

Pruitt KD, Tatusova T, Maglott DR (2007) NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 35: D61-65

Punglia RS, Burstein HJ, Winer EP, Weeks JC (2008) Pharmacogenomic variation of CYP2D6 and the choice of optimal adjuvant endocrine therapy for postmenopausal breast cancer: a modeling analysis. J Natl Cancer Inst. 100: 642-648

Purohit V (1998) Moderate alcohol consumption and estrogen levels in postmenopausal women: a review. Alcohol Clin Exp Res. 22: 994-997

Qu X, Zhang X, Qin A, et al (2013) Bone mineral density and risk of breast cancer in postmenopausal women. Breast Cancer Res Treat. 138: 261-271

Quail MA, Smith M, Coupland P, et al (2012) A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics 13: 341

Rae JM, Drury S, Hayes DF, et al (2012) CYP2D6 and UGT2B7 genotype and risk of recurrence in tamoxifen-treated breast cancer patients. J Natl Cancer Inst. 104: 452-460

Rahman N, Seal S, Thompson D, et al (2007) PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. Nature Genetics 39: 165-167

Raina D, Uchida Y, Kharbanda A, et al (2014) Targeting the muc1-c oncoprotein downregulates her2 activation and abrogates trastuzumab resistance in breast cancer cells. Oncogene 33: 3422-3431

Rakha EA, Boyce RW, Abd El-Rehim D, et al (2005) Expression of mucins (MUC1, MUC2, MUC3, MUC4, MUC5AC and MUC6) and their prognostic significance in human breast cancer. Mod Pathol. 18: 1295-1304

Ramensky V, Bork P, Sunyaev S (2002) Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 30: 3894-3900

Ramkaran P, Phulukdaree A, Khan S, et al (2015) Methylenetetrahydrofolate reductase C677T polymorphism is associated with increased risk of coronary artery disease in young South African Indians. Gene 571: 28-32

Ramsay DT, Kent JC, Hartmann RA, Hartman PE (2005) Anatomy of the lactating human breast redefined with ultrasound imaging. J Anat. 206: 525-534

Rankin SC (2000) MRI of the breast. Br J Radiol. 73: 806-818

Raymond E, Faivre S, Chaney S, et al (2002) Cellular and molecular pharmacology of oxaliplatin. Mol. Cancer Ther. 1: 227-235

Raza Ali H, Rueda OM, Chin S, et al (2014) Genome-driven integrated classification of breast cancer validated in over 7,500 samples. Genome Biology 15: 431

Reeves MD, Yawitch TM, Van der Merwe NC, et al (2004) BRCA1 mutations in South African breast and/or ovarian cancer families: evidence of a novel founder mutation in Afrikaner families. Int J Cancer 110: 677-682

Regan MM, Leyland-Jones B, Bouzyk M, et al (2012) CYP2D6 genotype and Tamoxifen response in postmenopausal women with endocrine-responsive breast cancer: the breast international group 1-98 trial. J Natl Cancer Inst 104: 441-451

Regitnig P, Reiner A, Dinges HP, et al (2002) Quality assurance for detection of estrogen and progesterone receptors by immunohistochemistry in Austrian pathology laboratories. Virchows Arch. 441: 328-334

Reinhardt H, Jiang H, Hemann M, Yaffe M (2009) Exploiting synthetic lethal interactions for targeted cancer therapy. Cell Cycle 8: 3112-3119

Renwick A, Thompson D, Seal S, et al (2003) ATM mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. Nat Genet. 38: 873-875

Reznikoff-Etiévan MF, Cayol V, Carbonne B ET, et al (2001) Factor V Leiden and G20210A prothrombin mutations are risk factors for very early recurrent miscarriage. BJOG 108: 1251-1254

Ricks-Santi LJ, Nie J, Marian C, et al (2013) BRCA1 Polymorphisms and Breast Cancer Epidemiology in the Western New York Exposures and Breast Cancer (WEB) Study. Genet. Epidemiol. 37: 504-511

Ring A and Dowsett M (2004) Mechanisms of tamoxifen resistance. Endocr Relat Cancer. 11: 643-658

Rios J and Puhalla S (2011) PARP inhibitors in breast cancer: BRCA and beyond. Oncology 25: 1014-1025

Ripperger T, Gadzicki D, Meindl A, Schlegelberger B (2008) Breast cancer susceptibility: current knowledge and implications for genetic counselling. Eur J Hum Genet. 17: 722-731

Rivenbark AG, O'Connor SM, Coleman WB (2013) Molecular and cellular heterogeneity in breast cancer: challenges for personalized medicine. Am J Pathol. 183: 1113-1124

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

Rohan TE, Jain MG, Howe GR, Miller AB (2000) Dietary folate consumption and Breast cancer risk. J Natl Cancer Inst 92: 266-269

Romond EH, Perez EA, Bryant J, et al (2005) Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. N Engl J Med. 353: 1673-1684

Room R, Babor T, Rehm J (2005) Alcohol and public health. The Lancet 365: 519-30

Rosato V, Bosetti C, Talamini R, et al (2011) Metabolic syndrome and the risk of breast cancer in postmenopausal women. Ann Oncol 22: 2687-2692

Rouleau M, Patel A, Hendzel MJ, et al (2010) PARP inhibition: PARP1 and beyond. Nat Rev Cancer 10: 293-301

Roy R, Chun J, Powell SN (2012) BRCA1 and BRCA2: important differences with common interests. Nature Reviews Cancer 12: 372

Royer-Pokora B, Kunkel L, Monaco A, et al (1986) Cloning the gene for the inherited disorder chronic granulomatous disease on the basis of its chromosomal location. Nature 322: 32-38

Ruark E, Snape K, Humburg P, et al (2013) Mosaic PPM1D mutations are associated with predisposition to breast and ovarian cancer. Nature 493: 406-410

Rummel S, Varner E, Shriver CD, Ellsworth RE (2013) Evaluation of BRCA1 mutations in an unselected patient population with triple-negative breast cancer. Breast Cancer Res Treat. 137: 119-125

Russo J and Russo IH (2006) The role of estrogen in the initiation of breast cancer. J Steroid Biochem Mol Biol. 102: 89-96

Saadat M (2012) Apolipoprotein E (APOE) Polymorphisms and Susceptibility to Breast Cancer: A Meta-Analysis. Cancer Res Treat. 44: 121-126

Saarenmaa I, Salminen T, Geiger U, et al (1999) The visibility of cancer on earlier mammograms in a population-based screening programme. Eur J Cancer 35: 1118-1122

Saleem A, Searle GE, Kenny LM, et al (2015) Lapatinib access into normal brain and brain metastases in patients with HER2 overexpressing breast cancer. EJNMMI Res. 5: 30

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors Proc Natl Acad Sci USA 74: 5463-5467

Sangrajrang S, Sato Y, Sakamoto H, et al (2010) Genetic polymorphisms in folate and alcohol metabolism and breast cancer risk: a case-control study in Thai women. Breast Cancer Res Treat 123: 885-893

Saniah AR and Zainal N (2010) Anxiety, depression and coping strategies in breast cancer patients on chemotherapy. Malaysian Journal of Psychiatry 19(2)

Sanyal AJ (2002) American Gastroenterological Association: AGA technical review on nonalcoholic fatty liver disease (national guidelines). Gastroenterology 123: 1705-1725

Sartor MA, Dolinoy DC, Jones TR, et al (2011) Genome-wide methylation and expression differences in HPV(+) and HPV(-) squamous cell carcinoma cell lines are consistent with divergent mechanisms of carcinogenesis. Epigenetics 6: 777-787

Sayin M and Bakkal BH (2011) Deep Venous Thrombosis in Breast Cancer Patients Using tamoxifen, a Hypothesis. Journal of Cancer Therapy 2: 607-609

Schaapveld M, Visser O, Louwman MJ, et al (2008) Risk of new primary nonbreast cancers after breast cancer treatment: a Dutch population-based study. J Clin Oncol 26*:* 1239-1246

Schierbeck LL, Rejnmark L, Tofteng CL, et al (2012) Effect of hormone replacement therapy on cardiovascular events in recently postmenopausal women: randomized trial. BMJ 345: e6409

Schneider JA, Rees DC, Liu YT, Clegg JB (1998) Worldwide distribution of a common methylenetetrahydrofolate reductase mutation. Am J Hum. Genet. 62: 1258-1260

Schoeman M, Apffelstaedt JP, Baatjes K, Urban M (2013) Implementation of a breast cancer genetic service in South Africa – lessons learned. South African Medical Journal 103: 529-533

Scholtz CL, Odendaal HJ, Thiart R, et al (2002) Analysis of two mutations in the MTHFR gene associated with mild hyperhomocysteinaemia: Heterogeneous distribution across ethnic groups. S Afr Med J 92: 464-467

Schrag D, Kuntz KM, Garber JE, Weeks JC (2000) Life expectancy gains from cancer prevention strategies for women with breast cancer and BRCA1 or BRCA2 mutations. JAMA 283: 617-624

Schroth W, Goetz MP, Hamann U, et al (2009) Association between CYP2D6 polymorphisms and outcomes among women with early stage breast cancer treated with tamoxifen. JAMA 302: 1429-1436

Schuster SC (2008) Next-generation sequencing transforms today's biology. Nat. Methods 5: 16-18

Schwahn B and Rozen R (2001) Polymorphisms in the methylenetetrahydrofolate reductase gene: clinical consequences. Am J Pharmacogenomics 1: 189-201

Schwarz JM, Rödelsperger C, Schuelke M, Seelow D (2010) MutationTaster evaluates disease-causing potential of sequence alterations. Nature Methods 7: 575-576

Scott SA (2001) Personalizing medicine with clinical pharmacogenetics. Genet Med. 13: 987-995

Seal S, Thompson D, Renwick A, et al (2006) Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. Nature Genetics 38: 1239-1241

Shang Y, Hu X, DiRenzo J, et al (2000) Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. Cell 103: 843-852

Shaw RJ and Cantley LC (2006) Ras, PI(3)K and mTOR signalling controls tumour cell growth. Nature 441: 424-430

Shen H, Spitz MR, Wang LE, et al (2001) Polymorphisms of methylene-tetrahydrofolate reductase and risk of lung cancer: a case-control study. Cancer Epidemiol Biomarkers Prev. 10: 397-401

Shen L, Kondo Y, Guo Y, et al (2007) Genome-wide profiling of DNA methylation reveals a class of normally methylated CpG island promoters. PLoS Genet. 3: 2023-2036

Shrubsole MJ, Shu XO, Ruan ZX, et al (2005) MTHFR genotypes and breast cancer survival after surgery and chemotherapy: a report from the Shanghai Breast Cancer Study. Breast Can Res Treat. 91: 73-79

Shukla PC, Singh KK, Quan A, et al (2011) BRCA1 is an essential regulator of heart function and survival following myocardial infarction. Nature Communications 2: 593

Shumaker SA, Legault C, Rapp SR, et al (2003) Estrogen Plus Progestin and the Incidence of Dementia and Mild Cognitive Impairment in Postmenopausal Women: The Women's Health Initiative Memory Study: A Randomized Controlled Trial. JAMA 289: 2651-2662

Siemianowicz K, Gminski J, Garczorz W, et al (2003). Methylenetetrahydrofolate reductase gene C677T and A1298C polymorphisms in patients with small cell and nonsmall cell lung cancer. Oncol Rep. 10: 1341-1344

Siepel A, Bejerano G, Pedersen JS, et al (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 15: 1034-1050

Sieri S, Chiodini P, Agnoli C et al (2014) Dietary Fat Intake and Development of Specific Breast Cancer Subtypes. JNCI J Natl Cancer Inst. doi: 10.1093/jnci/dju068

Silva PO and Gorini MIPC (2012) Validation of defining characteristics for the nursing diagnosis of fatigue in oncological patients. Rev. Latino-Am. Enfermagem 20: 504-510

Simpson ER (2001) Minireview: aromatase and the regulation of estrogen biosynthesis–some new perspectives. Endocrinology 142: 4589-4594

Siroy A, Abdul-Karim FW, Miedler J, et al (2013) MUC1 is expressed at high frequency in early-stage basal-like triple-negative breast cancer. Hum Pathol. 44: 2159-2166

Slamon DJ, Leyland-Jones B, Shak S, et al (2001) Use of Chemotherapy plus a Monoclonal Antibody against HER2 for Metastatic Breast Cancer That Overexpresses HER2. New England Journal of Medicine 344: 783-792

Slanger TE, Chang-Claude J, Wang-Gohrke S (2006) Manganese superoxide dismutase Ala-9Val polymorphism, environmental modifiers, and risk of breast cancer in a German population. Cancer causes and control 17: 1025-1031

Sledge GW, Loehrer PJ, Roth BJ, et al (1988) Cisplatin as first-line therapy for metastatic breast cancer. J Clin Oncol. 6: 1811-1814

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

Smith-Warner SA, Spiegelman D, Yaun SS, et al (1998) Alcohol and breast cancer in women: a pooled analysis of cohort studies. JAMA 279: 535-540

Snape K, Ruark E, Tarpey P, et al (2012) Predisposition gene identification in common cancers by exome sequencing: insights from familial breast cancer. Breast Cancer Res Treat. 134: 429-433

Sotiriou C, Neo SY, McShane LM, et al (2003) Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci USA 100: 10393-10398

Spector NL and Blackwell KL (2009) Understanding the mechanisms behind trastuzumab therapy for human epidermal growth factor receptor 2-positive breast cancer. Journal of Clinical Oncology 27: 5838-5847.

Spurdle AB, Whiley PJ, Thompson B, et al (2012) BRCA1 R1699Q variant displaying ambiguous functional abrogation confers intermediate breast and ovarian cancer risk. J Med Genet. 49: 525-532

Stephenson GD and Rose DP (2003) Breast cancer and obesity: an update. Nutr Cancer. 45: 1-16

Stevens R and Kalkwarf D (1990) Iron, radiation, and cancer. Environ Health Perspect. 87: 291-300

Stevenson JC (2006) HRT, osteoporosis and regulatory authorities Quis custodiet ipsos custodes? Hum Reprod. 21: 1668-1671

Stewart A, Bielajew C, Collins B (2006) A meta-analysis of the neuropsychological effects of adjuvant chemotherapy treatment in women treated for breast cancer. Clin Neuropsychol. 20: 76-89

Stover PJ (2004) Physiology of folate and vitamin B12 in health and disease. Nutr Rev. 62: S3-S12

Strachan T and Read AP (1999) Human molecular genetics (2nd edition). Wiley-Liss, New York
Stratton MR, Campbell PJ, Futreal PA (2009) The cancer genome. Nature. 458: 719-724

Su Y, Jiang Y, Sun S, et al (2015) Effects of HER2 genetic polymorphisms on its protein expression in breast cancer. Cancer Epidemiol. 39: 1123-1127

Su Y, Zheng Y, Zheng W, et al (2011) Distinct distribution and prognostic significance of molecular subtypes of breast cancer in Chinese women: A population-based cohort study. BMC Cancer 11: 292

Suppli NP, Deltour I, Damkjaer LH, et al (2011) Factors associated with the prescription of antidepressive medication to breast cancer patients. Acta Oncol 50: 243-251

Surekha D, Vishnupriya S, Sailaja K, et al (2008) Influence of apolipoprotein e gene polymorphism on the risk for breast cancer. Int J Hum Genet. 8: 277-282

Swen JJ, Wilting I, de Goede AL, et al (2008) Pharmacogenetics: From Bench to Byte. Clin Pharmacol Ther. 83: 781-787

Swift M, Reitnauer PJ, Morrell D, Chase CL (1987) Breast and other cancers in families with ataxia-telangiectasia. N Engl J Med. 316: 1289-1294

Swisher AK, Abraham J, Bonner D, et al (2015) Exercise and dietary advice intervention for survivors of triple-negative breast cancer: effects on body fat, physical function, quality of life, and adipokine profile. Support Care Cancer 23: 2995-3003

Tajouri L, Martin V, Gasparini C, et al (2006) Genetic investigation of methylenetetrahydrofolate reductase (MTHFR) and catechol-O-methyl transferase (COMT) in multiple sclerosis. Brain Res Bull. 69: 327-331

Tangutoori S, Baldwin P, Sridhar S (2015) PARP inhibitors: a new era of targeted therapy. Maturitas 81: 5-9

Taraseviciute A and Voelkel NF (2006) Severe pulmonary hypertension in postmenopausal obese women. European Journal of Medical Research 11: 198-202

Tell G, Fantini D, Quadrifoglio F (2010) Understanding different functions of mammalian AP endonuclease (APE1) as a promising tool for cancer treatment. Cell Mol Life Sci. 67: 3589-3608

Thakur S, Sarkar B1, Cholia RP, et al (2014) APE1/Ref-1 as an emerging therapeutic target for various human diseases: phytochemical modulation of its functions. Exp Mol Med. 46: e106

Thawnashom K, Tungtrongchitr R, Petmitr S, et al (2005) Methylenetetrahydrofolate reductase (MTHFR) polymorphism (C677T) in relation to homocysteine concentration in overweight and obese Thais. Southeast Asian J Trop Med Public Health 36: 2

The Cancer Association of South Africa. http://www.cansa.org.za. Accessed August 2015

The Prognosis of Breast Cancer Patients after Mastectomy and Immediate Breast Reconstruction: A Meta-Analysis

Thomas DB (1995) Alcohol as a cause of cancer. Environ Health Perspect 103:153-160

Thomas G, Jacobs KB, Kraft P, et al (2009) A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). Nature Genetics 41: 579-584

Thompson DS, Spanier CA, Vogel VG (1999) The Relationship Between Tamoxifen, Estrogen, and Depressive Symptoms. Breast J. 5: 375-382

Thomson TA, Hayes MM, Spinelli JJ, et al (2001) HER2/neu in breast cancer: interobserver variability and performance of immunohistochemistry with 4 antibodies compared with fluorescent in situ hybridization. Mod Pathol. 14: 1079-1086

Tian S, Roepman P, van't Veer LJ, et al (2010) Biological functions of the genes in the MammaPrint breast cancer profile reflect the hallmarks of cancer. Biomarker Insights 5: 129-138

Tiemeier H, van Tuijl HR, Hofman A, et al (2002) Vitamin B12, folate, and homocysteine in depression: the Rotterdam Study. Am J Psychiatry 59: 2099-2101

Tischkowitz M and Xia B (2010) PALB2/FANCN: Recombining cancer and Fanconi anemia. Cancer Research 70: 7353-7359

Toffoli G, Russo A, Innocenti F, et al (2003) Effect of methylenetetrahydrofolate reductase 677 C>T polymorphism on toxicity and homocysteine plasma level after chronic methotrexate treatment of ovarian cancer patients. Int J Cancer 103: 294-299

Toledo E, Salas-Salvadó J, Donat-Vargas C, et al (2015) Mediterranean Diet and Invasive Breast Cancer Risk Among Women at High Cardiovascular Risk in the PREDIMED Trial: A Randomized Clinical Trial. JAMA Intern Med. doi:10.1001/jamainternmed.2015.4838

Tomasello G1, Bedard PL, de Azambuja E, et al (2010) Brain metastases in HER2-positive breast cancer: the evolving role of lapatinib. Crit Rev Oncol Hematol. 75: 110-121

Tommiska J, Seal S, Renwick A, et al (2006) Evaluation of RAD50 in familial breast cancer predisposition. Int J Cancer 118: 2911-2916

Tong SY, Ha SY, Ki KD, et al (2009) The effects of obesity and HER2 polymorphisms as risk factors for endometrial cancer in Korean women. BJOG 116: 1046-1052

Treangen TJ and Salzberg SL (2012) Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nat Rev Genet 13: 36-46

Trentham-Dietz A, Newcomb PA, Nichols HB, Hampton JM (2007) Breast cancer risk factors and second primary malignancies among women with breast cancer. Breast Cancer Res Treat. 105: 195-207

Turkoz FP, Solak M, Petekkaya I, et al (2013) The prognostic impact of obesity on molecular subtypes of breast cancer in premenopausal women. J BUON 18: 335-341

Turner NC, Reis-Filho JS, Russell AM, et al (2007) BRCA1 dysfunction in sporadic basal-like breast cancer. Oncogene 26: 2126-2132

Ulrich CM (2007) Folate and cancer prevention: a closer look at a complex picture. Am J Clin Nutr 86: 271-273

Ulrich CM and Potter JD (2006) Folate supplementation: too much of a good thing? Cancer Epidemiol Biomarkers Prev 15: 189-193

Uniprot. http://www.uniprot.org. Accessed May 2015

Vahdaninia M, Omidvari S, Montazeri A (2010) What do predict anxiety and depression in breast cancer patients? A follow-up study. Social Psychiatry and Psychiatric Epidemiology 45: 355-361

Vahteristo P, Bartkova J, Eerola H, et al (2002) A CHEK2 genetic variant contributing to a substantial fraction of familial breast cancer. Am J Hum Genet. 71: 432-438

van den Brandt P, Spiegelman D, Yaun SS, et al (2000) Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. Am J Epidemiol. 152: 514-527

van der Merwe N, Bouwens CSH, Pienaar R, et al (2012) CYP2D6 genotyping and use of antidepressants in breast cancer patients: test development for clinical application. Metab Brain Dis. 27: 319-326

van der Merwe N, Peeters A, van Rensburg SJ, et al (2015) Implications of discordance between reference genomes applied in breast cancer exome sequencing. The Breast 2451: 87-150

van der Merwe NC and van Rensburg EJ (2009) Hereditary breast/ovariancancer and BRCA mutations: a South African perspective. Curr Oncol. 16: 347

van der Merwe NC, Hamel N, Schneider S-R et al (2012) A founder BRCA2 mutation in non-Afrikaner breast cancer patients of the Western Cape of South Africa. Clin Genet. 81: 179-184

van der Put NM, van der Molen EF, Kluijtmans LA, et al (1997) Sequence analysis of the coding region of human methionine synthase: relevance to hyperhomocysteinaemia in neural-tube defects and vascular disease. QJM 90: 511-517

van der Wall E, Donker TH, de Frankrijker E, et al (1993) Inhibition of the in vivo conversion of androstenedione to estrone by the aromatase inhibitor vorozole in healthy postmenopausal women. Cancer Res. 53: 4563-4566

Vardy J (2009) Cognitive function in breast cancer survivors. Cancer Treat Res 151: 387-419

Vehmanen L, Elomaa I, Blomqvist C, Saarto T (2006) Tamoxifen treatment after adjuvant chemotherapy has opposite effects on bone mineral density in premenopausal patients depending on menstrual status. J Clin Oncol. 24: 675-680

Vgontzas AN, Tan TL, Bixler OE, et al (1994) Sleep apnea and sleep disruption in obese patients. Archives of Internal Medicine 154: 1705-1711

Vinh-Hung V and Verschraegen C (2004) Breast-conserving surgery with or without radiotherapy: pooled-analysis for risks of ipsilateral breast tumour recurrence and mortality. J Natl Cancer Inst. 96: 115-121

Vogelstein B, Papadopoulos N, Velculescu VE (2013) Cancer genome landscapes. Science 339: 1546-1558

Vorobiof DA, Sitas F, Vorobiof G (2001) Breast cancer incidence in South Africa. J Clin Oncol. 19: 125S-127S

Wald DS, Law M, Morris JK (2002) Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. Br Med J. 325: 1202

Walsh T, Casadei S, Coats KH, et al (2006) Spectrum of mutations in BRCA1, BRCA2, CHEK2, and TP53 in families at high risk of breast cancer. JAMA 295: 1379-1388

Walsh T, Lee MK, Casadei S, et al (2010) Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing. Proc Natl Acad Sci. 107: 12629

Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research 38: e164

Wang LG and Chiao JW (2010) Prostate cancer chemopreventive activity of phenethyl isothiocyanate through epigenetic regulation (review). Int. J. Oncol. 37: 533-539

Warner E, Plewes DB, Hill KA, et al (2004) Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination. JAMA 292: 1317-1325

Warwick J, Pinney E, Warren RML, et al (2003) Breast density and breast cancer risk factors in a high-risk population. Breast 12: 10-16

Wegman P, Vainikka L, Stål O, et al (2005) Genotype of metabolic enzymes and the benefit of tamoxifen in postmenopausal breast cancer patients. Breast Cancer Res. 7: R284-R290

Weinhold B (2006) Epigenetics: The Science of Change. Environmental Health Perspectives 114: A160-A167

Weiwei Z, Liping C, Dequan L (2014) Association between dietary intake of folate, vitamin B6, B12 & MTHFR, MTR Genotype and breast cancer risk. Pak J Med Sci. 30: 106-110

Weston A and Godbold JH (1997) Polymorphisms of H-ras-1 and p53 in breast cancer and lung cancer: a meta-analysis. Environ. Health Perspect. 105: 919-926

Wettergren Y, Odin E, Carlsson G, Gustavsson B (2010) MTHFR, MTR, and MTRR polymorphisms in relation to p16INK4A hypermethylation in mucosa of patients with colorectal cancer. Mol Med. 16: 425-432

Whitehead J, Carlile T, Kopecky KJ (1985) Wolfe mammographic parenchymal patterns. A study of the masking hypothesis of Egan and Mosteller. Cancer 56: 1280-1286

Whitworth P, Stork-Sloots L, de Snoo FA, et al (2014) Chemosensitivity Predicted by BluePrint 80-Gene Functional Subtype and MammaPrint in the Prospective Neoadjuvant Breast Registry Symphony Trial (NBRST). Ann Surg Oncol. 21: 3261-3267

WHO and FAO report (2001) Human Vitamin and Mineral Requirements. Chapter 4: pp.59

Wilfond BS, Rothenberg KH, Thomson EJ, et al (1997) Cancer genetic susceptibility testing: Ethical and policy implications for future research and clinical practice. J Law Med Ethics 25: 243-251

Wilson A, Platt R, Wu Q, et al (1999) A common variant in methionine synthase reductase combined with low cobalamin (vitamn B12) increases risk for spina bifida. Mol. Genet. Metab. 67: 317-323

Wilson DM, Kim D, Berquist BR, Sigurdson AJ (2011) Variation in base excision repair capacity. Mutat Res 711: 100-112

Wilson PWF, D'Agostino RB, Sullivan L, et al (2002) Overweight and Obesity as Determinants of Cardiovascular Risk: The Framingham Experience. Arch Intern Med. 162: 1867-1872

Winer EP (2005) Optimizing endocrine therapy for breast cancer. J Clin Oncol. 23: 1609-1610

Winer EP, Hudis C, Burstein HJ, et al (2004) American Society of Clinical Oncology technology assessment on the use of aromatase inhibitors as adjuvant therapy for postmenopausal women with hormone receptor-positive breast cancer: Status Report. J Clin Oncol. 23: 619-629

Wirapati P, Sotiriou C, Kunkel S, et al (2008) Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. Breast Cancer Res 10: R65

Wishart DS (2015) Is Cancer a Genetic Disease or a Metabolic Disease? EBioMedicine 2: 478-479

Wolk A, Gridley G, Svensson M, et al (2001) A prospective study of obesity and cancer risk (Sweden). Cancer Causes Control. 12: 13-21

Wooster R and Weber BL (2003) Breast and ovarian cancer. N Engl J Med. 348: 2339-2347

World Health Organization (2008) The Global Burden of Disease: 2004 Update. Geneva: World Health Organization.

Wright GE, Niehaus DJ, Drögemöller BI, et al (2010) Elucidation of CYP2D6 genetic diversity in a unique African population: implications for the future application of pharmacogenetics in the Xhosa population. Ann Hum Genet. 74: 340-350

Wu YL, Ding XX, Sun YH et al. (2013) Association between MTHFRC677T polymorphism and depression: an updated meta-analysis of 26 studies. Prog Neuropsychopharmacol Biol Psychiatry 46: 78-85

Xanthoudakis S, Miao G, Wang F, et al (1992) Redox activation of Fos-Jun DNA binding activity is mediated by a DNA repair enzyme. EMBO J 11: 3323-3335

Xia W, Mullin RJ, Keith BR, et al (2002) Anti-tumour activity of GW572016: a dual tyrosine kinase inhibitor blocks EGF activation of EGFR/erbB2 and downstream Erk1/2 and AKT pathways. Oncogene 21: 6255-6263

Yager JD and Davidson NE (2006) Estrogen carcinogenesis in breast cancer. New Engl J Med 354: 270-282

Yamada K, Chen Z, Rozen R, Matthews RG (2001) Effects of common polymorphisms on the properties of recombinant human methylenetetrahydrofolate reductase. Proc Natl Acad Sci USA 98: 14853-14858

Yamaji T, Iwasaki M, Sasazuki S, et al (2009) Methionine synthase A2756G polymorphism interacts with alcohol and folate intake to influence the risk of colorectal adenoma. Cancer Epidemiol. Biomarkers Prev. 18: 267-274

Yamnik RL and Holz MK (2010) mTOR/S6K1 and MAPK/RSK signaling pathways coordinately regulate estrogen receptor alpha serine 167 phosphorylation. FEBS Lett. 584: 124-128

Yamnik RL, Digilova A, Davis DC, et al (2009) S6 kinase 1 regulates estrogen receptor alpha in control of breast cancer cell proliferation. J Biol Chem. 284: 6361-6369

Yang M, Rong T, Huang Z, et al (2005) Operable breast cancer: A clinical analysis of 6,263 cases. Chinese Journal of Clinical Oncology 2: 761-766

Yeo Z, Wong JCL, Rozen SG, et al (2014) Evaluation and optimisation of indel detection workflows for ion torrent sequencing of the BRCA1 and BRCA2 genes. BMC Genomics 15: 516

Yersal O and Barutca S (2014) Biological subtypes of breast cancer: Prognostic and therapeutic implications. World Journal of Clinical Oncology 5: 412-424

Yin M, Liao Z, Liu Z, et al (2011) Functional polymorphisms of base excision repair genes XRCC1 and APEX1 predict risk of radiation pneumonitis in patients with non-small cell lung cancer treated with definitive radiation therapy. Int J Radiat Oncol Biol Phys. 81: e67-73

Yood MU, Wells KE, Alford SH, et al (2012) Cardiovascular outcomes in women with advanced breast cancer exposed to chemotherapy. Pharmaco epidemiol Drug Saf. 21: 818-827

Zanger UM, Raimundo S, Eichelbaum M (2004) Cytochrome P450 2D6: overview and update on pharmacology, genetics, biochemistry. Naunyn Schmiedebergs Arch Pharmacol. 369: 23-37

Zanger UM, Turpeinen M, Klein K, Schwab M (2008) Functional pharmacogenetics/genomics of human cytochromes P450 involved in drug biotransformation. Anal Bioanal Chem. 392: 1093-1108

Zghair AN, Sharma R, Sharma AK (2014) Hormone responsive breast cancer and BRCA1 mutation: mechanism, regulation and iron-mediated effects. Curr Pharm Biotechnol. 15: 1113-1124

Zhang X, Wei J, Zhou L, et al (2013) A functional BRCA1 coding sequence genetic variant contributes to risk of esophageal squamous cell carcinoma. Carcinogenesis 34: 2309-2313

Zhang Y, He BS, Pan YQ, et al. (2001) Association of OGG1 Ser326Cys polymorphism with colorectal cancer risk: a meta-analysis. Int J Colorectal Dis. 26: 1525-1530

Zhang Y, Iwata T, Yamamoto J, et al (2011) FTIR study of light-dependent activation and DNA repair processes of (6-4) photolyase. Biochemistry 50: 3591-3598

Zhao W, Hu L, Xu J, et al (2013) Polymorphisms in the base excision repair pathway modulate prognosis of platinum-based chemotherapy in advanced non-small cell lung cancer. Cancer Chemother Pharmacol. 71: 1287-1295

Zheng W, Long J, Gao YT, et al (2009) Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nature Genetics 41: 324-328

Zhou Q, Yin W, Du Y, Lu J (2014) For or against adjuvant trastuzumab for pT1a-bN0M0 breast cancer patients with HER2-positive tumours: a meta-analysis of published literatures. PLoS One 9: e83646

Zijno A, Andreoli C, Leopardi P, et al (2003) Folate status, metabolic genotype, and biomarkers of genotoxicity in healthy subjects. Carcinogenesis. 24: 1097-1103

Zineh I and Pacanowsi MA (2011) Pharmacogenomics in the assessment of therapeutic risks versus benefits: inside the United States Food and Drug Administration. Pharmacotherapy 31: 729-735

Zunarelli E, Nicoll JA, Migaldi M, Trentini GP (2000) Apolipoprotein E polymorphism and breast carcinoma: correlation with cell proliferation indices and clinical outcome. Breast Cancer Res Treat. 63: 193-198

Zurrida S, Bassi F, Arnone P, et al (2011) The Changing Face of Mastectomy (from Mutilation to Aid to Breast Reconstruction). Int J Surg Oncol. 2011: 980158

CHAPTER 8

APPENDICES

APPENDIX I

Appendix I describes the DNA extraction protocol used to extact DNA from whole blood, followed by the procedure for quantification of extracted DNA using the NanoDrop® ND-1000 spectrophotometer.

DNA extraction Protocol

The QIAGEN QIAamp® DNA Blood Maxi Kit DNA extraction protocol (spin protocol) for whole blood:

1. The DNA extraction procedure started with dispensing 500 µl QIAGEN Protease stock solution into the bottom of a 50 ml centrifuge tube. This solution contains the enzyme (protease) responsible for lysing or breaking up the cells, expelling their cellular components and releasing their DNA into the solution.
2. 5 ml of blood sample was added to the stock solution and very briefly mixed (to create homogenous solution for adequate lysis), using a vortex machine.
3. 6 ml buffer AL was added to the tube and mixed by inverting the tube 15 times, followed by vigorous shaking (using a vortex machine) for approximately 1 minute.
4. The homogenized solution was incubated at 70°C in a dry block for 15 minutes. After incubation, 5ml ethanol (96% -100%) was added to the sample and mixed (to ensure binding of the lysate to the membrane) by inverting the tube 10 times followed by vigorous shaking.
5. The solution was carefully transfered onto the QIAamp® Maxi column (membrane) inside a 50 ml centrifuge tube and was centrifuged at 3000 rpm for 3 minutes.
6. The QIAamp® Maxi column was transferred to a clean 50 ml centrifuge tube and the filtrate discarded to prevent the nozzle of the QIAamp® Maxi column from being submerged in the filtrate (reduces the washing efficacy).
7. 5 ml Buffer AW1 was added to the QIAamp® Maxi column and centrifuged at 5000 rpm for 1 minute. This step was repeated for buffer AW2 but centrifuged for 15 minutes.
8. The QIAamp® Maxi column was placed into a clean 50 ml centrifuge tube and the collection tube containing the filtrate was discarded.
9. 600 µl distilled water equilibrated to room temperature (15-25°C) was added directly onto the membrane of the QIAamp® Maxi column.
10. This was incubated at room temperature for 5 minutes, followed by centrifugation at 5000 rpm for 2 minutes.
11. The same volume of distilled water was added and this step repeated, but centrifuged for 5 minutes. The eluted solution now contained the newly extracted DNA.

Spectrophotometry

Subsequent to switching on the instrument and selecting the Nucleic Acid analysis option, the instrument pedestal was cleaned and Nuclease Free Water added onto the pedestal to initialize the instrument. RNase Free water (2 µl) was placed on the pedestal to calibrate the instrument before taking any measurements. Upon completion of the blanking process, all the samples were measured and stored. The pedestal was cleaned between each measurement with 70% ethanol followed by double distilled water. The ratio absorbance reading at 230 nm, 260 nm and 280 nm was used to assess the purity of the DNA. Purity values for all samples were within the accepted range (260/280 absorbance ratio: 1.6-1.9; 260/230 absorbance ratio : >1.9) indicating the absence of contaminants such as proteins, salts or phenols in the sample.

Validation of prioritized variants by the Pathology Research Facility research team

Whole exome sequencing identified potential causative gene variations and functional polymorphisms in the breast cancer patients selected for extended mutation analysis using the newly developed exome prescreen algorithm (EPA). Conventional polymerase chain reaction (PCR) was performed followed by direct DNA sequencing of the resulting PCR fragments and/or and/or real-time PCR for validation of WES results. Results obtained for the APEX1 and MYH9 variants are provided as examples of common and rare variants with no clinical relevance in the patients studied. Confirmation of the RAD50 and MUC1 variants provide examples of rare missense mutations identified in the index patient with familial breast cancer.

Identification of wild type, heterozygous and homozygous genotypes for the APEX1 and MYH9 variants are listed in Table 8.2, followed by high-throughput genotyping using ABI™ TaqMan® SNP technology with standard reaction mixes and cycling parameters (described in the subjects and methods Tables 5.4 and 5.5).

8.1. Conventional PCR and Sanger sequencing results

The specific amplicons obtained for the APEX1 and MYH9 variants using conventional primers visualized with ethidium Bromide in a 2 % agarose gel is presented in figure 8.1.

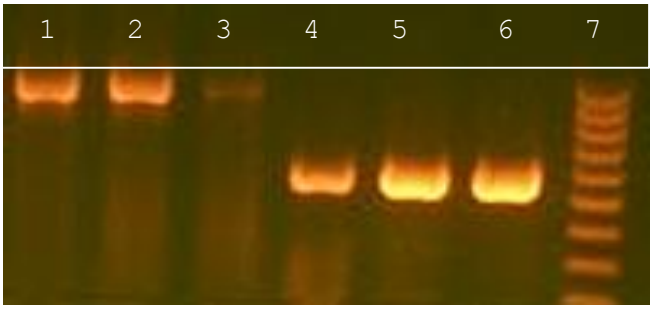


Figure 8.1: A 2% (w/v) agarose gel visualized with 0.0001% (v/v) ethidium bromide (EtBr). Following PCR amplification using APEX1 D148E and MYH9 G236G primer sets, lanes 1-3 contain amplicons of 898bp and lanes 4-6 contain 547bp. Lane 7 contains a 100bp ladder.

Figure 8.2 depicts the Sanger sequencing results of APEX1 D148E and MYH9 G236G, performed by the Central Analytical Facility (CAF) of Stellenbosch University. The sequencing results of each genotype in relation to the risk-associated allele when not detected (wild type), heterozygous or homozygous are presented in electropherograms for the exome-sequenced index and unrelated breast cancer patients in figures 8.2 (APEX1-wild type), 8.3 (APEX1-heterozygous), 8.4 (APEX1-homozygous), 8.5 (MYH9- wild type) and 8.6 (MYH9-heterozygous). Figures 8.7 to 8.9 relates to the sequences of the index patient (MUC1-heterozygous), her mother (MUC1-wild type) and the unrelated breast cancer patient (MUC1-wild type), respectively. The index patient's mother's sequencing results are omitted since she possesses the same APEX1 and MYH9 genotypes as her daughter. No homozygous genotype was detected for the rare MYH9 variant. Although forward (sense) and reverse (anti-sense) sequencing were carried out for each genotype to demonstrate the analytical validity, only the forward sequencing results supported by the reverse sequences are shown here.

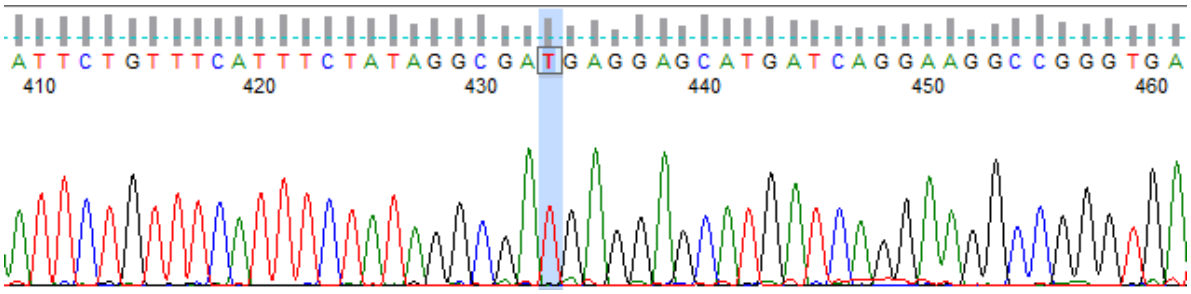


Figure 8.2. Electropherogram illustrating the forward sequencing reaction of an amplified PCR product (index breast cancer patient) obtained with the APEX1 D148E primer set. The nucleotide position of the SNP is indicated by the highlighted region and corresponds to a genotype of TT, representing a wild type genotype.

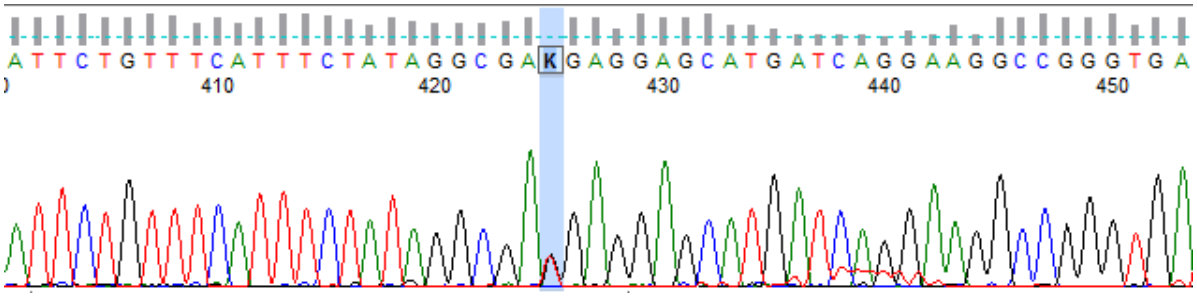


Figure 8.3. Electropherogram illustrating the forward sequencing reaction of an amplified PCR product (non-exome sequenced breast cancer patient) obtained with the APEX1 D148E primer set. The nucleotide position of the SNP is indicated by the highlighted region and corresponds to a genotype of GT, representing a heterozygous genotype.

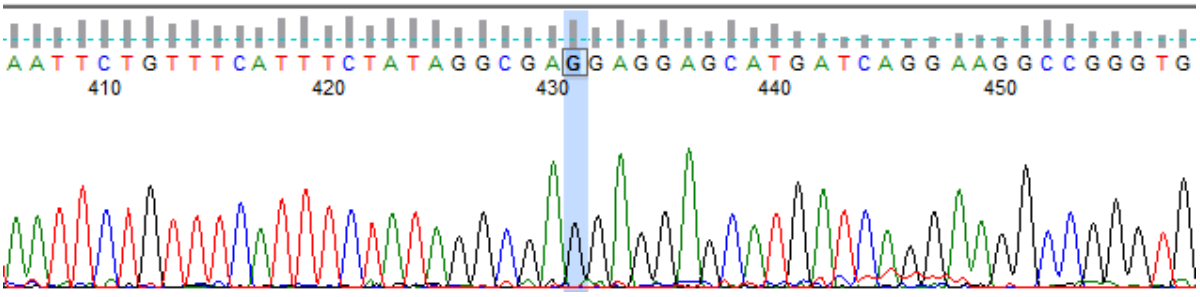


Figure 8.4. Electropherogram illustrating the forward sequencing reaction of an amplified PCR product (unrelated breast cancer patient) obtained with the APEX1 D148E primer set. The nucleotide position of the SNP is indicated by the highlighted region and corresponds to a genotype of GG, representing a homozygous mutant genotype.

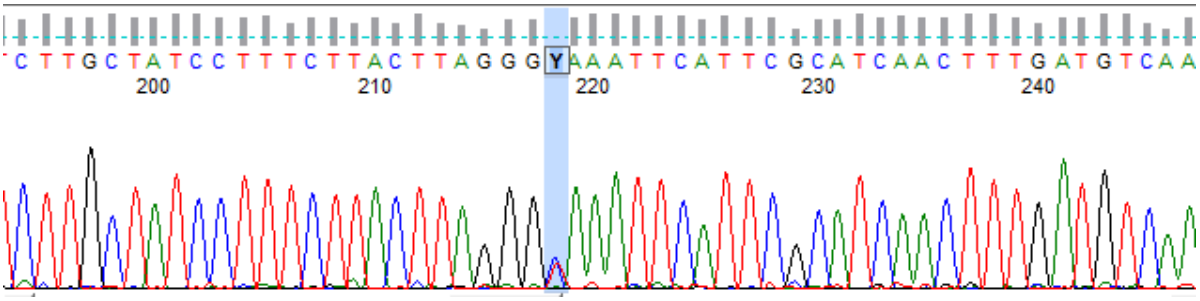


Figure 8.5. Electropherogram illustrating the forward sequencing reaction of an amplified PCR product (index patient) obtained with the MYH9 G236G primer set. The nucleotide position of the variant is indicated by the highlighted region and corresponds to a genotype of CT, representing a heterozygous genotype.

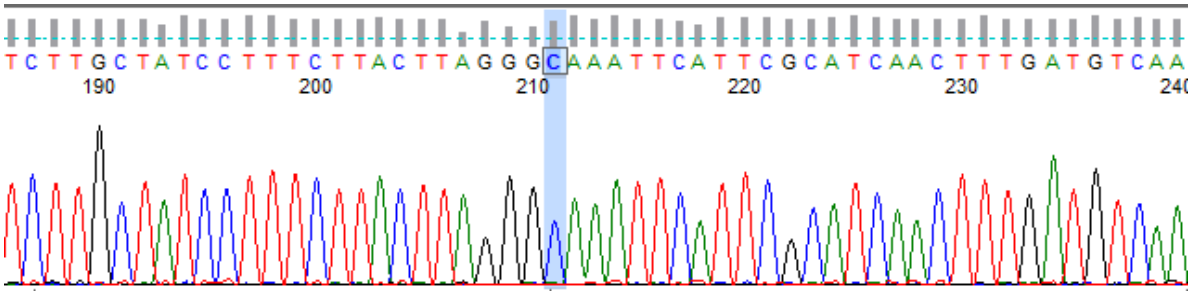


Figure 8.6. Electropherogram illustrating the forward sequencing reaction of an amplified PCR product obtained with the MYH9 G236G primer set. The nucleotide position of the variant is indicated by the highlighted region and corresponds to a genotype of CC, representing a wild type genotype detected in the exome-sequenced unrelated breast cancer patient and extended breast cancer patient sample.

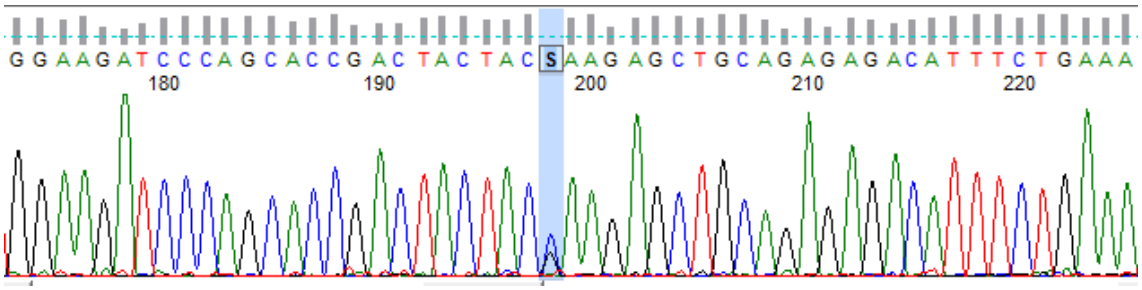


Figure 8.7. Electropherogram illustrating the forward sequencing reaction of an amplified PCR product obtained with the MUC1 Q67E primer set. The nucleotide position of the variant is indicated by the highlighted region and corresponds to a genotype of CG, representing a heterozygous genotype detected in the exome-sequenced index breast cancer patient.

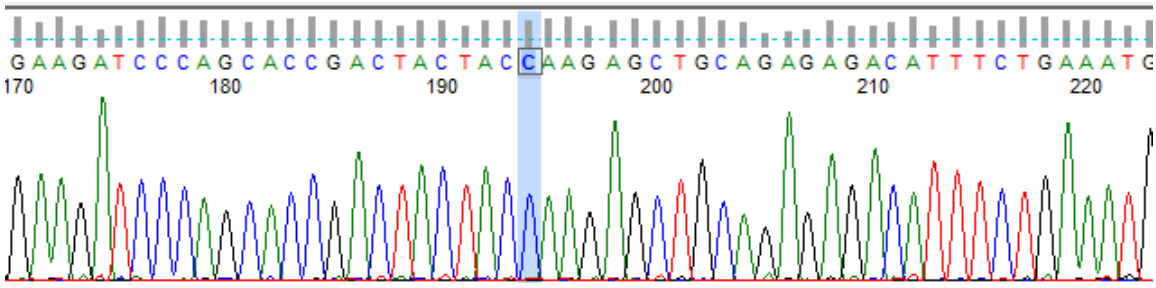


Figure 8.8. Electropherogram illustrating the forward sequencing reaction of an amplified PCR product obtained with the MUC1 Q67E primer set. The nucleotide position of the variant is indicated by the highlighted region and corresponds to a genotype of CC, representing a wild type genotype detected in the index patient's mother with breast cancer.

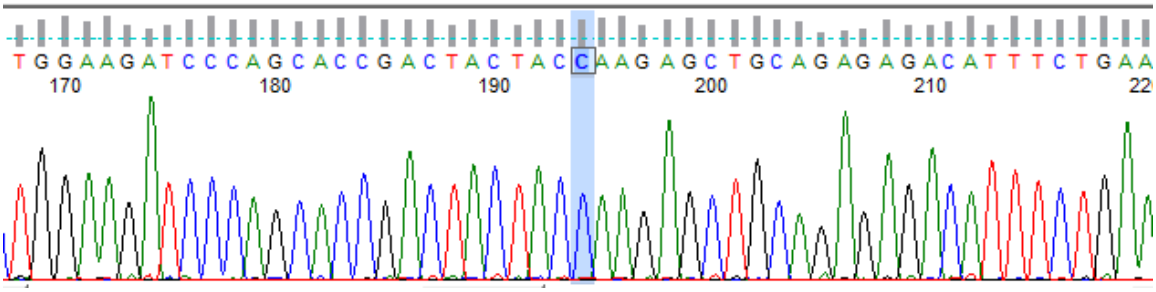


Figure 8.9. Electropherogram illustrating the forward sequencing reaction of an amplified PCR product obtained with the MUC1 Q67E primer set. The nucleotide position of the variant is indicated by the highlighted region and corresponds to a genotype of CC, representing a wild type genotype detected in the unrelated breast cancer patient.

The above mentioned samples with known genotypes verified by Sanger sequening were subsequently used as positive controls to screen the extended patient and control groups using real-time PCR methodology.

8.2. Real-time PCR results

The SNP assays included in the CVD multi-gene test (subjects and methods Table 4.7) were successfully performed on the Corbett Rotor-Gene™ 6000 series multiplexing system and the WES-identified variants on the Roche LightCycler® 480 II real-time(RT)-PCR machine, by the Pathology Research Facility (PRF). Real-time PCR experiments yielded clear amplification as well as precise genotype calling of samples. This result was further demonstrated by the verified

exome-sequenced samples as well as internal controls (K1-K3) included in each real-time PCR experiment and loaded on each plate. The non-template controls (NTC's) revealed clear clustering per plate with no indication of fluorescence and therefore amplification. This confirmed and depicted the absence of contamination during the experimental procedure. Samples with dubious genotyping results were repeated whereas failed samples that displayed little or no fluorescence, attributed to poor DNA quality (degredation or fragmentation of the DNA), were excluded from subsequent analyses.

Due to the bulk of genotyping results generated from this study only one sample batch of each assay is illustrated in this section to demonstrate the results obtained with the use of the Corbett Rotor-Gene or Roche LightCycler machines. Verification and analytical validation of gene variants identified by WES concluded the final phase of this study.

The allelic discrimination analysis for the RAD50 (1153 C>T, rs139372231) assay is presented in Figure 8.10 with the legend and genotypes illustrated by Table 8.1. The lines without intermitted circles represent the (C) allele, which is detected by FAM™-labelled probes, while the lines with the intermitted circles signifies the homozygote (T) allele detected by the VIC®-labelled probes. Amplification of both alleles simultaneously indicative of significant fluorescence of both dyes represents the heterozygous (C/T) genotype. A threshold level (0.25) was established as a discriminatory parameter during analysis to indicate the presence of allele specific fluorescence above the threshold setting.

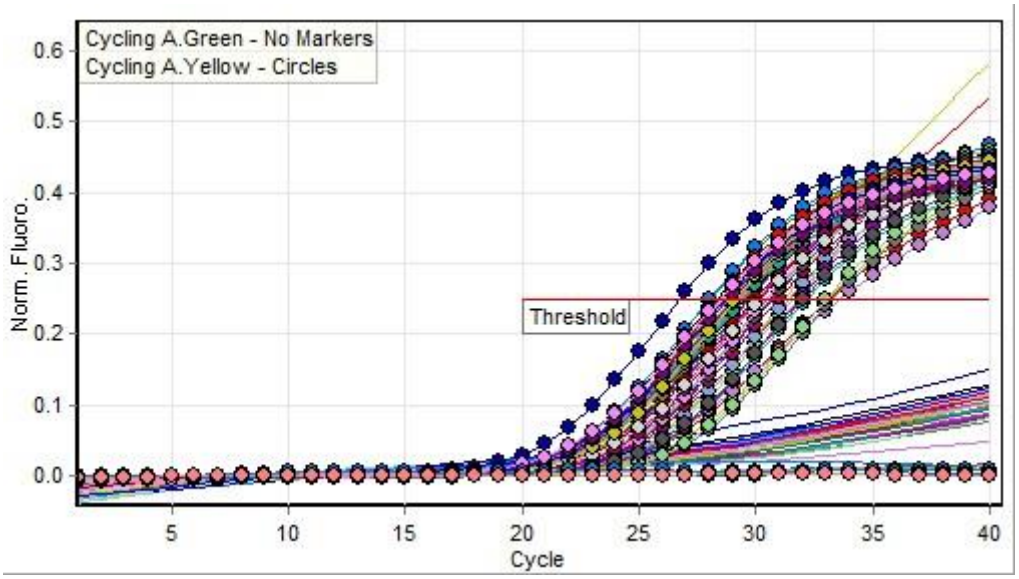


Figure 8.10. Allelic discrimination analysis of RAD50 (1153 C>T, rs139372231) using the ABI™ TaqMan® (C_171053490_10) genotyping assay. (Normalized fluorescence vs. number of cycles) performed on the Corbett Rotor-Gene™ 6000.

Table 8.1. Legend for figure 8.6, specifying genotypes of the samples based on Allelic discrimination data and Scatterplot analysis on the Corbett Rotor-Gene™ 6000.

No.	Colour	Name	Genotype	Cycling A.Green	Cycling A.Yellow
1		Index patient	Heterozygous	Reaction	Reaction
2		Mother	Heterozygous	Reaction	Reaction
3		Unrelated patient	Wild Type	No Reaction	Reaction
4		Sample 1	Wild Type	No Reaction	Reaction
5		Sample 2	Wild Type	No Reaction	Reaction
6		Sample 3	Wild Type	No Reaction	Reaction
7		Sample 4	Wild Type	No Reaction	Reaction
8		Sample 5	Wild Type	No Reaction	Reaction
9		Sample 6	Wild Type	No Reaction	Reaction
10		Sample 7	Wild Type	No Reaction	Reaction
11		Sample 8	Wild Type	No Reaction	Reaction
12		Sample 9	Wild Type	No Reaction	Reaction
13		Sample 10	Wild Type	No Reaction	Reaction
14		Sample 11	Wild Type	No Reaction	Reaction
15		Sample 12	Wild Type	No Reaction	Reaction
16		Sample 13	Wild Type	No Reaction	Reaction
17		Sample 14	Wild Type	No Reaction	Reaction
18		Sample 15	Wild Type	No Reaction	Reaction
19		Sample 16	Wild Type	No Reaction	Reaction

Table 8.1. Legend for figure 8.6, specifying genotypes of the samples based on Allelic discrimination data and Scatterplot analysis on the Corbett Rotor-Gene™ 6000.

No.	Colour	Name	Genotype	Cycling A.Green	Cycling A.Yellow
20		Sample 17	Wild Type	No Reaction	Reaction
21		Sample 18	Wild Type	No Reaction	Reaction
22		Sample 19	Wild Type	No Reaction	Reaction
23		Sample 20	Wild Type	No Reaction	Reaction
24		Sample 21	Wild Type	No Reaction	Reaction
25		Sample 22	Wild Type	No Reaction	Reaction
26		Sample 23	Wild Type	No Reaction	Reaction
27		Sample 24	Wild Type	No Reaction	Reaction
28		Sample 25	Wild Type	No Reaction	Reaction
29		Sample 26	Wild Type	No Reaction	Reaction
30		Sample 27	Wild Type	No Reaction	Reaction
31		Sample 28	Wild Type	No Reaction	Reaction
32		Sample29	Wild Type	No Reaction	Reaction
33		Sample 30	Wild Type	No Reaction	Reaction
34		K1	Wild Type	No Reaction	Reaction
35		K2	Wild Type	No Reaction	Reaction
36		NTC1	Wild Type	No Reaction	Reaction
37		NTC2	Wild Type	No Reaction	Reaction
38		NTC3	Wild Type	No Reaction	Reaction

Generated by the Roche LightCycler® 480 II software, the allele-specific PCR (ASPCR) amplification curve for the MUC1 Q67E assay is presented in Figure 8.11. The purple curve represents the ancestral allele while the blue curve represents the variant allele. In Figure 8.11, both curves showed amplification in the index patient's sample. The curves that display early amplification (i.e low cycle number) represent the ancestral or wild type allele, while the variant allele amplifies during later cycles (i.e high cycle number). Amongst the exome-sequenced individuals, only the index patient was confirmed to be heterozygous for MUC1 Q67E.

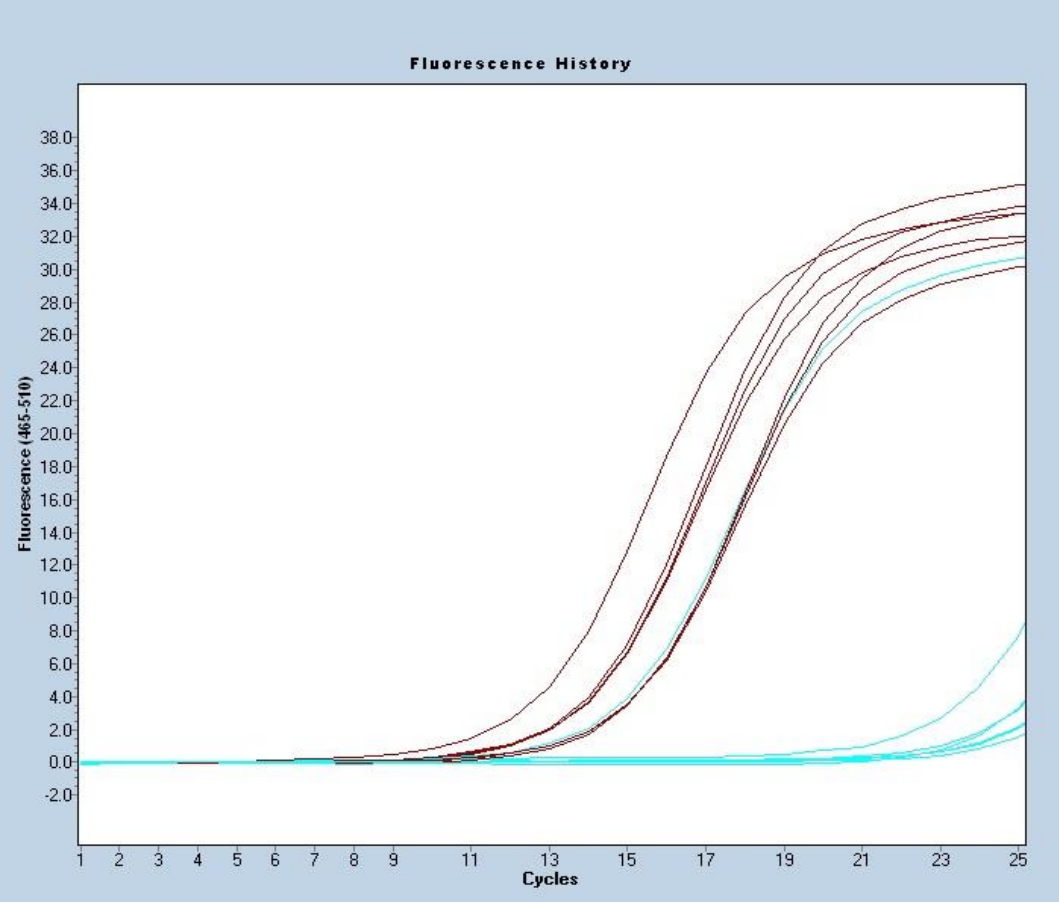


Figure 8.11. ASPCR amplification curve of MUC1 (Q67E) depicting fluorescence vs. number of cycles, performed on the Roche LightCycler® 480 II.

In addition to endpoint genotyping allelic discrimination of APEX1 D148E on the Roche LightCycler® 480 II, scatterplot analysis was also performed to further validate the genotyping results. Scatterplot analysis was applied using two channels simultaneously. Genotype acquisition is based on specific regions found on the scatterplot as well as relative expression of amplification of the green (FAM™ = Y: G-allele) and yellow (VIC® = X: T-allele) channels

respectively. Upon normalization, the different fold increases of each channel and log transformation accentuated the differences in expression of each individual sample.

The scatterplot analysis and allelic discrimination plot for the APEX1 D148E assay is presented in Figures 8.12 and 8.13, respectively, with the legend and genotypes depicted in Table 8.2. Genotype determinations were based on the dual VIC® and FAM™ Minor Groove Binding (MGB) probes. Table 8.3 shows the comparison of genotype distribution allele frequencies for APEX1 rs1130409 between breast cancer patients and controls who participated in the chronic disease screen.

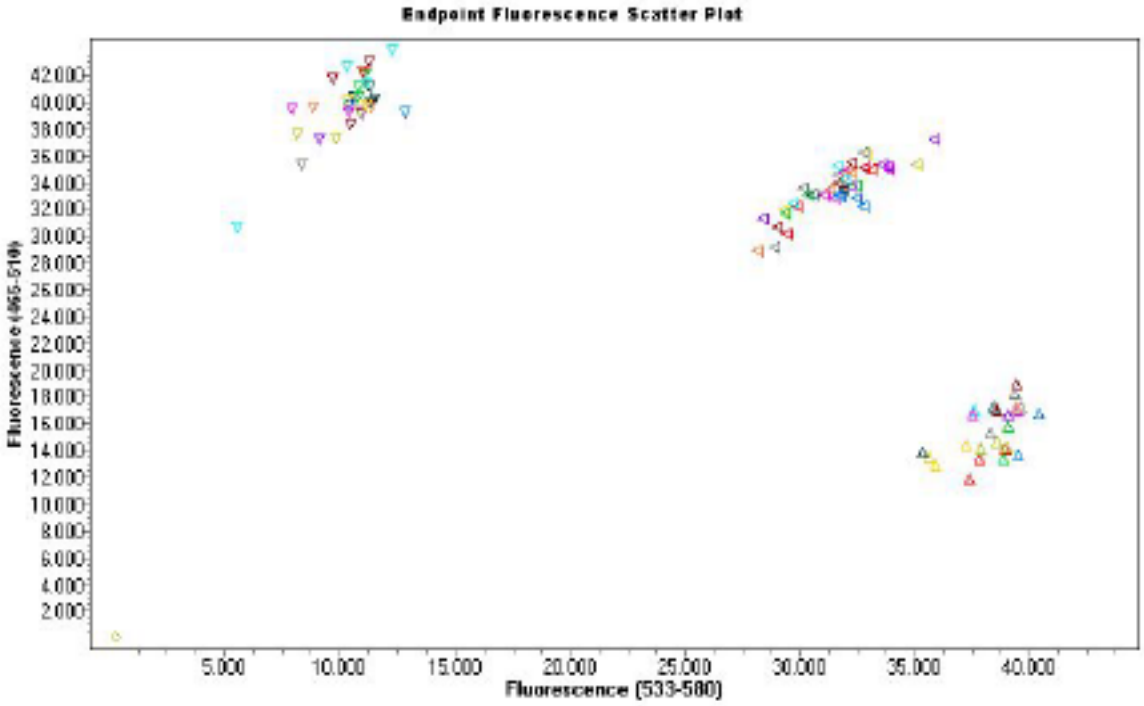


Figure 8.12. Genotypes grouped by scatterplot analysis (FAM™ fluorescence vs. VIC® fluorescence) of the ABI™ TaqMan® (C_8921503_10) APEX1 D148E assay. [Allele Y (G- FAM labelled) vs Allele X (T- VIC labelled)]. Upper left cluster = Homozygous GG, middle cluster = Heterozygous GT, bottom left cluster = Homozygous TT. The legend is presented in Table 8.2.

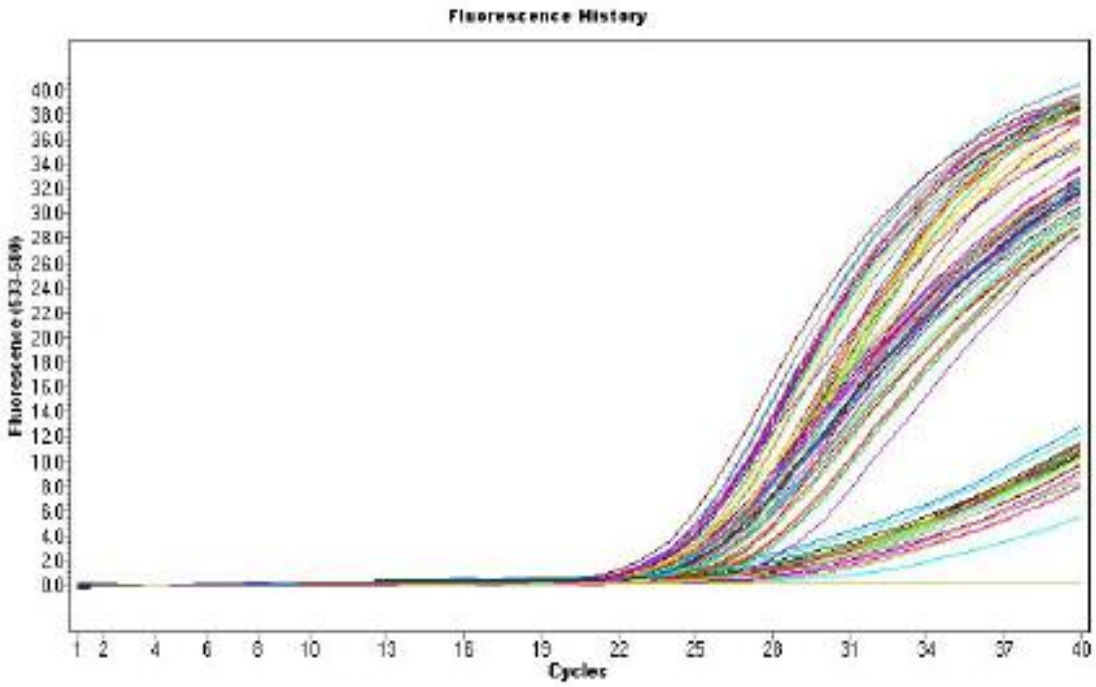


Figure 8.13. Allelic discrimination analysis (fluorescence vs. number of cycles) of APEX1 D148E using the ABI™ *Taq*Man® (C_7514879_10) genotyping assay.

Table 8.2. Legend for Figures 8.8 and 8.9, specifying genotypes of the samples based on Endpoint analysis data and Scatterplot analysis.

APEX1_breast cancer patient and controls (Endpoint Genotyping)

Results

Inc	Pos.	Sample Name	Endpoint Fluorescence Allele X	Endpoint Fluorescence Allele Y	Call	Score	Status
☑	A1	Sample 1	12.85	39.26	Allele Y	0.85	
☑	A2	Index patient	37.85	13.38	Allele X	0.91	
☑	A3	Sample 3	32.49	33.75	Both Alleles	0.99	
☑	A4	Sample 4	31.54	32.75	Both Alleles	0.99	
☑	A5	Unrelated patient	8.34	35.29	Allele Y	0.96	
☑	A6	Index patient mother	35.64	13.47	Allele X	0.95	
☑	A7	Sample 7	29.03	30.64	Both Alleles	0.99	
☑	A8	Sample 8	5.55	30.72	Allele Y	0.87	
☑	A9	Sample 9	35.36	13.96	Allele X	0.99	
☑	A10	Sample 10	28.22	28.94	Both Alleles	0.98	
☑	A11	Sample 11	39.10	16.71	Allele X	0.95	
☑	A12	Sample 12	37.94	14.10	Allele X	0.94	
☑	B1	Sample 13	10.40	39.72	Allele Y	0.99	
☑	B2	Sample 14	29.44	30.21	Both Alleles	0.98	
☑	B3	Sample 15	39.14	15.78	Allele X	1.00	
☑	B4	Sample 16	31.05	33.07	Both Alleles	0.98	
☑	B5	Sample 17	28.93	29.20	Both Alleles	0.96	
☑	B6	Sample 18	29.31	31.87	Both Alleles	0.96	
☑	B7	Sample 19	10.48	38.34	Allele Y	0.97	
☑	B8	Sample 20	37.62	17.10	Allele X	0.88	
☑	B9	Sample 21	11.54	40.21	Allele Y	0.94	
☑	B10	Sample 22	31.32	33.42	Both Alleles	0.98	
☑	B11	Sample 23	10.95	39.08	Allele Y	0.95	
☑	B12	Sample 24	38.64	14.70	Allele X	0.96	
☑	C1	Sample 25	40.48	16.78	Allele X	0.97	

Table 8.3. Genotype distribution comparisons of APEX1 rs1130409 between breast cancer patients and controls who participated in the chronic disease screen.

Genotypes	Coloured breast cancer patients	Caucasian breast cancer patients	Coloureds vs Cauasians	Caucasian Controls	Caucasian Pts vs Controls
APEX1 T>G	n=82	n=81		n=147	
TT	23 (28%)	26 (32%)		32 (22%)	
GT	42 (51%)	33 (41%)	0.35	78 (53%)	0.23
GG	17 (21%)	22 (27%)		37 (25%)	
T	88 (54%)	85 (52%)	0.89	142 (48%)	1.0
G	76 (46%)	77 (48%)		152 (52%)	

Table 8.4. List of variants identified in three whole exome-sequenced breast cancer patients for verification and validation using Sanger Sequencing and Taqman genotyping.

Gene ID	dbSNP ID#	Nulceotide Change	Primer ID	Sequence	Length (bp)	GC Content (%)	PrimerBLAST Tm (°C)	Amplicon (bp)	TaqMan Assay ID
APEX1	rs1130409	g.6865T>G c.444T>G	APEX1-FWD	TGCTGCTTGACTCGAACTCC	20	55.00	60.32	898	C___8921503_10
			APEX1-REV	TGCAGGTAACAGAGAGTGGG	20	55.00	59.02		
MYH9	rs754650591	g.71200C>T c.708C>T	MYH9-FWD	GTGTGGGATCAGCTTGTCTC	20	55.00	58.27	547	N/A
			MYH9-REV	TAGACGGCTTTCAACAGCTC	20	50.00	57.92		
RAD50	rs139372231	g.36865C>T c.1153C>T	RAD50-FWD	ATCCACATGCTCAGGGGTAC	20	55.00	59.16	528	C_171053490_10
			RAD50-REV	GCCAAAATGGAGTCCAACC	19	52.60	56.78		
MUC1	N/A	Q67E	MUC1-FWD	ATTCCCAGCCACCACTCTGA	20	55	60.84	493	N/A
			MUC1-REV	CCCAACCTTAAGTGCACCAGT	21	52.38	60.48		

APPENDIX II

JOURNAL PUBLICATIONS

1. **van der Merwe N**, Peeters AV, van Rensburg SJ, Pienaar FM, Bezuidenhout J, Kotze MJ. Implications of discordance between reference genomes applied in breast cancer exome sequencing. Breast 2015; 24 (Suppl 1): S121.

2. Kotze MJ, Lückhoff HK, Peeters AV, Baatjes K, Schoemann M, van der Merwe L, Grant KA, Fisher LR, **van der Merwe N**, Pretorius J, van Velden DP, Myburgh EJ, Pienaar FM, van Rensburg SJ, Yako YY, September AV, Moremi KE, Cronje FJ, Tiffin N, Bouwens CSH, Bezuidenhout J, Apffelstaedt JP, Hough FS, Erasmus RT, Schneider JW. Genomic medicine and risk prediction across the disease spectrum. Critical Reviews in Clinical Laboratory Sciences 2015, 19: 1-18.

3. Delport D, Schoeman R, **van der Merwe N**, van der Merwe L, Fisher LR, Geiger DH, Kotze, M. Significance of dietary folate intake, homocysteine levels and MTHFR 677 C>T genotyping in South African patients diagnosed with depression: test development for clinical application. Metabolic Brain Disease 2014; 29: 377-384.

4. **van der Merwe N**, Pienaar R, van Rensburg SJ, Bezuidenhout J, Kotze MJ. Identification of breast cancer patients at increased risk of “chemobrain”: Case study and review of the literature. South African Journal of Psychiatry 2013; 19 (3): 125.

5. **van der Merwe N**, Bouwens CHS, Pienaar FM, van der Merwe L, Yako YY, Geiger DH, Kotze MJ. CYP2D6 genotyping and use of antidepressants in breast cancer patients: test development for clinical application. Metabolic Brain Disease 2012; 27: 319-326.