Investigation of the role of vitamin D metabolism in South African breast cancer patients using a pathology-supported genetic testing platform

by

Abisola Oyedele Okunola

Dissertation presented in fulfilment of the requirements for the degree of Doctor of Philosophy (Chemical Pathology) in the Faculty of Medicine and Health Sciences Stellenbosch University

Supervisor

Professor Maritha J Kotze

Co-supervisors

Professor Rajiv T Erasmus

Professor Annalise E Zemlin

Dr Rispah Torrorey-Sawe

March 2020

DECLARATION

I the undersigned, hereby declare that the work contained in this thesis is my original work and that I have not previously submitted it, in its entirety or in part at any other University for a Degree.

Signature: _____ Date: _____

By submitting this dissertation 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.

ABSTRACT

The high global breast cancer incidence drives the development of novel genomic approaches for disease prevention and targeted treatment. Towards this goal, a pathology-supported genetic testing (PSGT) platform was established to facilitate risk management of non-communicable diseases across the continuum of care, ranging from early-stage to metastatic disease and cancer survivorship. The causes and consequences of low vitamin D levels recently reported in the majority of postmenopausal breast cancer patients treated with aromatase inhibitors at the Tygerberg Academic Hospital, in the Western Cape Province of South Africa, were addressed in this study using PSGT to translate genomic findings into clinical practice.

The aim was to determine the relationship between clinical characteristics, tumour histopathology and genetic variation underlying vitamin D metabolism in postmenopausal breast cancer patients at increased risk of osteoporosis, identified as a significant co-morbidity in the study population.

Clinical and lifestyle information of 116 postmenopausal women with known vitamin D status diagnosed with breast carcinoma between 2014 and 2017 was extracted from a central genomics database linked to a biobank of DNA samples extracted from blood. Whole exome sequencing (WES) was performed on the Ion Torrent platform, followed by variant calling of vitamin D-related genes, while simultaneously assessing *BRCA1/2* mutation status. Allele-specific real-time polymerase chain reaction (PCR), Sanger sequencing and long-range nanopore sequencing using the pocket-size MinION device were used to verify the WES results and to screen for variants in the vitamin D receptor (*VDR)* and E-cadherin (*CDH1)* genes beyond the coding regions covered by WES.

Seasonal variation (p = 0.009) and high body mass index (BMI) (p = 0.032) contributed significantly to vitamin D levels, with the lowest values recorded during winter. WES initially performed in 10 breast cancer patients selected based on vitamin D levels at extreme upper and lower ranges, identified *GC* rs4588 (c.1364C>A, T455K) as a potential contributing factor

to vitamin D deficiency in the five patients with ultra-low vitamin D levels (≤12 ng/mL). However, 2/5 patients with levels in the upper extreme of vitamin D (>30 ng/mL), also tested positive for this variant and no significant association was detected after extended genotyping in 100 South African patients using real-time PCR. Sanger sequencing subsequently performed in 14 breast cancer patients diagnosed with osteoporosis prior to initiation of aromatase inhibitor therapy, highlighted the potential significance of genetic variation in the *VDR* gene. WES analysis of *VDR* in an extended sample of 55 breast cancer patients furthermore confirmed the significant effect of genetic variation in this gene on bone health ($p < 0.001$).

The *CDH1* gene known to be activated by VDR was furthermore analysed in patients stratified by tumour type. *CDH1* c.G671A (p.R224H) detected in a breast cancer patient with invasive carcinoma of no special type (ICNST) was classified as benign, since pathogenic germline *CDH1* variants are associated with invasive lobular carcinoma and diffuse gastric cancer, but not ICNST. *CDH1* c.A1298G (p.D433G) was detected in a patient with invasive lobular carcinoma together with a pathogenic *BRCA1* variant detected by WES. Although this finding supports a likely benign classification for *CDH1* p.D433G as reported in the international ClinVar database, a family history of stomach cancer raised the possibility of a *CDH1* modifier gene effect on *BRCA1* gene expression.

New insights gained through integration of pathology and genomic findings were incorporated into a pharmaco-diagnostic algorithm applicable to hormone receptor-positive postmenopausal breast cancer patients. The PSGT platform facilitated interpretation of research results of study participants through use of WES and recommendation of genetic counselling where appropriate.

OPSOMMING

Die hoë globale voorkoms van borskanker dryf die ontwikkeling van nuwe genomiese benaderings tot siekte voorkoming and geteikende behandeling. Om hierdie doel te verwesenlik is ‘n patologie-ondersteunde genetiese toets platform (POGT) daargestel om risikohantering oor die kontinuum van sorg te fasiliteer, wat strek van vroeë tot metastatiese siekte en oorlewing van kanker. Die oorsaak en gevolge van lae vitamien D vlakke wat onlangs gerapporteer is in die meerderheid van postmenopousale borskanker pasiënte wat met aromatase inhibitore behandel word by die Tygerberg Akademiese Hospitaal, in die Wes Kaap provinsie van Suid-Afrika, is aangespreek in hierdie studie deur POGT te gebruik om genomiese bevindings om te skakel in kliniese praktyk.

Die doel was om die verband te bepaal tussen kliniese karaktertrekke, tumor histopatologie en genetiese variasie onderliggend aan vitamien D tekort in postmenopousale borskanker pasiënte met verhoogde risiko vir osteoporose, wat as ‘n belangrike ko-morbiditeit geïdentifiseer is in die studiepopulasie.

Kliniese en leefstyl inligting van 116 postmenopousale vroue met bekende vitamien D status, nuut gediagnoseer met borskanker tussen 2014 and 2017, is selekteer uit ‘n sentrale genomiese databasis wat gekoppel is aan ‘n biobank van DNA monsters wat geëkstraheer is uit bloed. Heel eksoom volgordebepaling (HEV) is uitgevoer op die Ion Torrent platform, gevolg deur variant selektering van vitamien D-verwante gene, met gelyktydige bepaling van *BRCA1/2* mutasie status. Alleel-spesifieke reël-tyd polymerase ketting reaksie (PKR), Sanger DNA volgordebepaling, en lang-fragment nanopore volgordebepaling is uitgevoer met gebruik van ‘n sak-grootte MinION apparaat om die HEV uitslae te bevestig en te sif vir variante in die vitamien D reseptor (*VDR*) en E-cadherin *(CDH1)* gene buite die kodering areas wat deur HEV gedek word.

Seisoenale variasie (p = 0.009) en hoë liggaam gewig indeks (p = 0.032) het statisties betekenisvol bygedra tot vitamien D vlakke, met die laagste waardes tydens winter gemeet.

HEV wat aanvanklik gedoen is in 10 geselekteerde borskanker pasiënte gebaseer op vitamien D wat in the uiterste van die hoogste en laagste vlakke gemeet het, het *GC* rs4588 (c.1364C>A, T455K) geïdentifiseer as 'n potensieël bydraende faktor tot vitamien D tekort in vyf pasiënte met ultra-laag vitamien D vlakke (≤12 ng/mL). Twee uit vyf pasiënte met die hoogste vlakke van vitamin D (>30 ng/mL) het ook postief getoets vir hierdie *GC* variant, en 'n uitgebreide genotipering deur gebruik van reël-tyd PKR in 100 Suid-Afrikaanse pasiënte nie 'n statisties betekenisvolle assosiasie met vitamien D vlakke getoon het nie. Sanger volgordebepaling wat is daarna uitgevoer in 14 borskanker pasiënte gediagnoseer met osteoporose voor die aanvang van aromatase inhibitor behandeling, het die kliniese relevansie van genetiese variasie in die *VDR* geen beklemtoon.

HEV analise van *VDR* in 'n uitgebreide aantal van 55 borskanker pasiënte het die betekenisvolle effek van *VDR* genetiese variasie op beengesondheid verder bevestig ($p < 0.001$). Die *CDH1* geen wat deur VDR geaktiveer word is verder ondersoek in pasiënte wat gestratisifeer is volgens tumor tipe. Die *CDH1* c.G671A (p.R224H) variant wat waargeneem is in 'n borskanker pasiënt met infiltrerende karsinoom van geen spesiale tipe is geklassifiseer as benigne, aangesien patogeniese kiemlyn *CDH1* variante geassosieer word met infiltrerende lobulêre karsinoom (ILK) en diffuse maagkanker, en nie infiltrerende karsinoom van geen spesiale tipe nie. *CDH1* c.A1298G (p.D433G) is in 'n pasiënt met ILK tesame met 'n patogeniese *BRCA1* variant waargeneem met gebruik van HEV. Alhoewel hierdie bevinding 'n waarskynlik benigne klassifikasie ondersteun soos gerappporteer in die internasionale ClinVar databasis, het die familiegeskiedenis van maagkanker die moontlikheid van 'n *CDH1* modifiseerder geen effek op *BRCA1* geen uitdrukking uitgelig.

Nuwe insigte soos verkry deur die integrasie van patologie en genomiese bevindinge in hierdie studie, is geïnkorporeer in n farmako-diagnostiese algoritme, toepaslik tot hormoon-positiewe postmenopousale borskanker pasiënte. Die POGT platform fasiliteer interpretasie van navorsingresultate van studie deelnemers deur die gebruik van HEV asook aanbeveling van genetiese raadgewing waar nodig.

DEDICATION

I dedicate this PhD thesis to Almighty God for giving me the grace to endure and persevere through this journey. Indeed, nothing worthwhile is ever easy.

ACKNOWLEDGEMENT

My profound appreciation to my supervisors and co-supervisors for their immense contributions to the successful completion of this study. To Professor Maritha Kotze, for accepting me to undertake this study in her research group. I am grateful for investing in me to attend workshops and conferences, at both local and international levels. Thanks so much for your time, criticisms, patience and understanding most especially during the write up of this dissertation. It was a privilege to have been under your tutelage. I am amazed by your genuine and strong desire to build a solid foundation for the incorporation of translational genetic studies to routine tests to aid better treatment on the African continent. You are indeed a woman with kind heart. Thanks to my co-supervisors Professor Erasmus, Professor Zemlin and Dr Torrorey-Sawe for the contributions of your expertise into this study. Thank you for the opportunity given me to express my mind to you even on issues not related to my studies. I learnt from Professor Erasmus “to always aim high and pursue your dreams no matter what the condition may present to you”, from Professor Zemlin “to persevere because it will end all well if you do not give up” and from Dr Torrorey-Sawe “not to look down on yourself”. My appreciation also goes to Dr Karin Baatjes for allowing me to use her database for this study. Thanks for this kind-heartedness.

A special thank you to Ms Sonja Krige, Chemical Pathology Divisional Secretary. You were never tired of my complaints and enquiries. Many thanks to staff and students in the Pathology Research Facility Laboratory: To Dr Armand Peeters, for those genetics lectures, laboratory and bioinformatics trainings; Ms Janine Cronje, for your kindness; Ms Kelebogile Moremi, I appreciate your laboratory training and encouragement; Dr Nicole Van der Merwe, for proof-reading the thesis draft; Ms Welile Dube, for editing and proof-reading some parts of the thesis draft. All your patience and time invested in this context is highly appreciated.

I appreciate Dr Jyothi Chabillal, the Faculty Doctoral Officer. My first personal meeting with you in June 2016 and your continued support contributed immensely to the completion of my study. Thanks for providing shoulders that postgraduate students in the Faculty can lean on. Thanks to Professor Martin Kidd for verifying the results of the statistical analysis performed.

I also acknowledge the South African Medical Research Council (SAMRC), Cancer Association of South Africa (CANSA), Technology Innovation Agency (TIA), the Philip Sceales and Janet Antrobus Breast Cancer Research Trust and Stellenbosch University for providing funding, including travel grants for this study. Appreciation is also expressed to the Wellcome Genome Campus, Hinxton, United Kingdom for the Next Generation Sequencing Workshop. Experience gained from the workshop contributed to the successful use of nanopore sequencing in this study.

Many thanks to my parents, Barrister and Mrs Okunola as well as Professor and Mrs Ojo for their encouragement, prayers, moral and financial support, despite being far away from home. Thanks so much for your advice to stay strong and remain focused when it really got tough. I also appreciate my brothers and sister: Dr ‘Seye and ‘Deji Okunola as well as Mrs ‘Joju Fawole and her family for their financial and moral support. Dr ‘Mike Ojo for his financial support and assistance in editing the thesis. ‘Tosin Ojo for your support and care.

This acknowledgement will not be complete without saying a special and wonderful thank you to my wife, Olutoyin, for her enduring support through this very hard journey. To my children, Eriifeoluwa “big sister” who joined the family at the beginning of the study. Also to Oluwatimileyin, whose arrival came during the writing phase of this study. I am trusting God to uphold and perfect all things concerning you. Both of you were a source of strength for me to forge ahead throughout this study.

My friends, ‘Dekunle Adedeji, ‘Kunle Odegbemi and ‘Segun Atanda are also thanked for their calls and messages to check on us. Indeed, you are friends.

Above all, thanks to God for placing all these people in my path through this journey.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ASIP	Agouti-signalling protein
ATM	Ataxia-telangiectasia mutated gene
ATP	Adenosine triphosphate
BAM	Binary alignment map
BLOC	Biogenesis of lysosomal organelles complex
BMI	Body mass index
BRCA	Breast cancer gene
BRIP	Breast cancer interacting protein C terminal helicase
CAF	Central Analytical Facility
CDH1	Cadherin 1
CHEK	Checkpoint kinase
CNR	Cannabinoid receptor
CVD	Cardiovascular diseases
CYP1A1	Cytochrome P450 family 1 subfamily A member 1
CYP17	Cytochrome P450 family 17
CYP19A1	Cytochrome P450 family 19 subfamily A member 1
CYP2R1	Cytochrome P450 family 2 subfamily R member 1
CYP24A1	Cytochrome P450 family 24 subfamily A member 1
CYP27B1	Cytochrome P450 family 27 subfamily B member 1
DBP	Vitamin D binding protein

DCIS	Ductal carcinoma *in situ*
DCT	Dopachrome tautomerase
DHCR7	7-dehydrocholesterol reductase
DKK	Dickkopf-related protein
dsDNA	Double-stranded deoxyribonucleic acid
DNTP	Deoxynucleotide triphosphate
DTNBP	Dystrobrevin binding protein 1
EDN	Endothelin
EDTA	Ethylenediaminetetraacetic acid
ER	Estrogen receptor
EXOC2	Exocyst complex component 2
FSH	Follicle stimulating hormone
GC	Group-specific component
gDNA	Genomic deoxyribonucleic acid
GLOBOCAN	Global cancer observatory
GQS	Genomic quality score
GSTs	Glutathione transferases
GWAS	Genome-wide association study
HER2	Human epidermal growth factor receptor 2
HPLC	High-performance liquid chromatography
HRT	Hormone replacement therapy

HWE	Hardy-Weinberg equilibrium
ICNST	Invasive carcinoma of no special type
IHC	Immunohistochemistry
ILC	Invasive lobular carcinoma
IRF	Interferon regulatory factor
LCIS	Lobular carcinoma in situ
LC-MS/MS	Liquid chromatography-mass spectrometry
MAF	Minor allele frequency
MITF	Melanocyte inducing transcription factor
MIR	Micro ribonuclease
MLPH	Melanophin
mRNA	Messenger ribonucleic acid
MTAP	Methylthioadenosine phosphorylase
mVDR	membrane vitamin D receptor
MYO	Myosin
NADSYN	Glutamine-dependent synthase
NAT	N-acetyltransferase
NGS	Next Generation Sequencing
NCDs	Non communicable diseases
NHLS	National Health Laboratory Service
OCA	Oculocutaneous albinism

PALB	Partner and localizer of breast cancer gene
PAX	Paired box gene
PCR	Polymerase chain reaction
PLDN	Pallidin
PMEL	Premelanosome protein
POC	Point of care
PolyPhen	Polymorphism phenotyping
PR	Progesterone receptor
PRKACG	Protein kinase cytosine adenosine monophosphate-activated catalytic subunit gamma
PSGT	Pathology-supported genetic testing
PTEN	Phosphatase and tensin homologue
RAB	Ras-related protein
RACK	Receptor for activated C kinase
RAD50	Double-stranded break repair protein
RAD51C	Rad 51 paralog C
RAS	Rat sarcoma
RCT	Randomised controlled trial
RNA	Ribonucleic acid
SNPs	Single nucleotide polymorphisms
SNV	Single nucleotide variant

SIFT	Sorting intolerant from tolerant
SOX	(Sex determining region)-box
STK	Serine/threonine kinase
TNBC	Triple negative breast cancer
TNF-α	Tumour necrosis factor alpha
TP53	Tumour protein 53
TVC	Torrent variant caller
TYRP1	Tyrosinase-related protein 1
UVB	Ultraviolet B
UK	United Kingdom
VCF	Variant call format file
VDR	Vitamin D receptor gene
VUS	Variant of uncertain clinical significance
WES	Whole exome sequencing
WHO	World Health Organization

LIST OF FIGURES

LIST OF TABLES

OUTLINE OF THE DISSERTATION

Chapter 1 includes the general introduction, research aim and objectives as well as the layout of the thesis. **Chapter 2** is a comprehensive literature review of the research topic including an evaluation of vitamin D biochemistry and the genes involved in this metabolic pathway that affects expression of VDR. Risk factors, classification and subtypes of breast cancer are furthermore discussed in the context of prevention, diagnosis and treatment. Finally, the application of next generation sequencing technologies and nanopore long-range sequencing are also explained in this context. **Chapter 3** provides a detailed report on the materials and methods used, with referral to the ethics approval process followed during the course of this study. **Chapter 4** covers a descriptive analysis of the study population based on vitamin D levels determined at presentation for genotype-phenotype association studies. Whole exome sequencing results are compared between two different sequencing platforms: Ion Torrent and Oxford Nanopore MinION. **Chapter 5** is the discussion of the results in relation to current knowledge, in order to highlight the novel contribution from the research performed. The incorporation of new knowledge into the pathology-supported genetic testing platform is explained in **Chapter 6** as the conclusion on the impact of the study.

CHAPTER 1

INTRODUCTION

According to the 2018 World Health Organization (WHO) statistics, 41 million deaths were recorded globally in 2016, of which approximately 70% were attributed to non-communicable diseases (NCDs). Most deaths were due to four major NCDs: cancers, cardiovascular disease (CVD), diabetes and chronic respiratory diseases (Bray et al., 2018). Breast cancer is the first among the 6 leading types of cancer in women worldwide (GLOBOCAN 2018). **Figure 1.1** shows the distribution of the six leading cancer types among females worldwide.

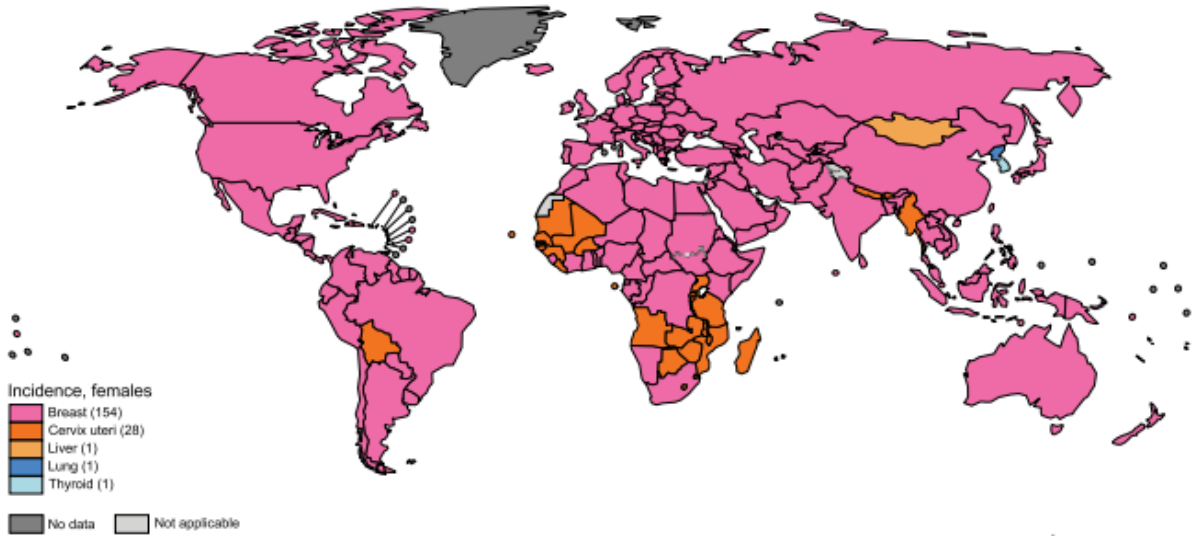


Figure 1.1: Worldwide geographical distribution of common types of cancer among females reported in 2018. Source: Global Cancer Statistics 2018 (GLOBOCAN 2018). Global Cancer Observatory (https://gco.iarc.fr/today). Permission from: Data source: GLOBOCAN 2018

The mortality rate when compared with incidence rate of breast cancer is higher on the African and Asian continents than in America and Europe (Ferlay et al., 2018). The breakdown of this analysis is shown in the **Figure 1.2**

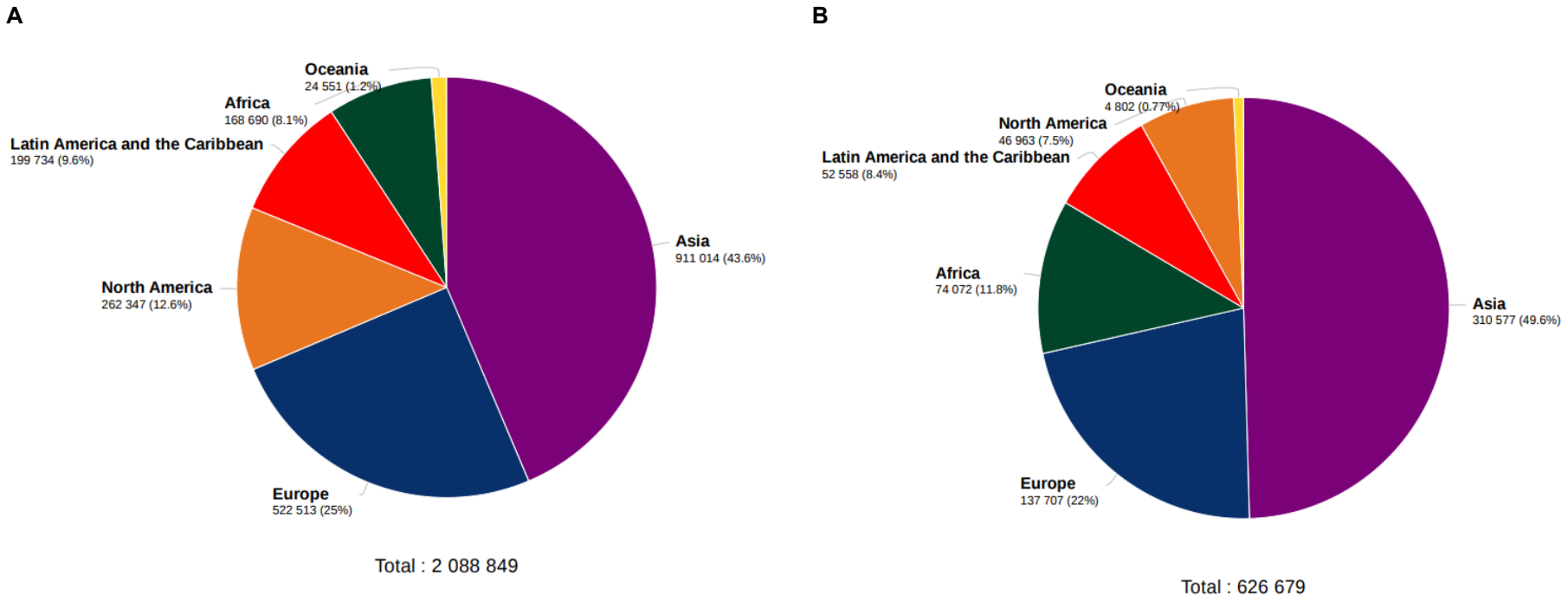


Figure 1.2: Worldwide estimated incidence (A) and mortality (B) rates for female breast cancer reported in 2018. The different areas of the pie chart reflect the proportion of the total incidence or mortality in the different continents. Permission from: Data source: GLOBOCAN 2018. Graph production: Global Cancer Observatory (https://gco.iarc.fr/today)

The increasing mortality rate of breast cancer in developing countries has been attributed to improved socioeconomic circumstances, involving movement from rural to urban settlements in search of higher education. This may be accompanied by adoption of a sedentary lifestyle with adverse effects on health and well-being (Bray et al., 2018).

The high global incidence of cancer drives the development of novel genomic approaches for prevention and treatment in an affordable yet efficient manner. *BRCA1* and *BRCA2*, categorized among the high penetrance cancer susceptibility genes are the most frequently mutated tumour suppressor genes in familial breast cancer (Mersch et al., 2015). Although carriers of pathogenic *BRCA1/2* germline variants have a significantly increased lifetime risk for breast cancer (60-85%) (Mehrgou & Akouchekian, 2016), incomplete mutation penetrance indicates the involvement of other genetic and environmental factors as modifiers of disease risk and metastatic potential (Cooper et al., 2013). These may include vitamin D deficiency and variation in the vitamin D receptor (*VDR*) gene, associated with malignancy and activation of mechanisms underlying drug resistance (Graziano et al., 2016). The functional relationship detected between the vitamin D/VDR axis and deficiencies in DNA repair factors in senescent cells may contribute to genomic instability, allowing senescence bypass and tumorigenesis. Strategies to optimize vitamin D status are therefore important, especially in African countries where cancer is increasing (Moukayed & Grant, 2013).

The importance of optimal vitamin D status in the prevention of osteoporosis and bone fractures is undisputed, while the effect on cancer risk remains uncertain. The reason why randomized controlled trials (RCTs) may fail to demonstrate causal connections with vitamin D levels in cancer patients has partly been ascribed to ignorance of critical biological criteria during study design (Lappe & Heaney, 2012). Individual genetic differences in vitamin D metabolism affecting treatment response are not adequately understood or quantified at present. RCTs of pharmaceutical drugs are performed with the assumption that the trial is the only source of the agent tested. Vitamin D levels however, are influenced by additional factors including age, gender, genetics and multiple environmental factors. These include latitude,

seasonal changes, body mass index (BMI), physical activity, smoking, diet and alcohol consumption (Shi et al., 2014). The relatively high prevalence of vitamin D deficiency among breast cancer patients with low physical activity and smokers led to recommendations for supplementation and other lifestyle modifications that may improve vitamin D status in breast cancer patients.

The favourable effect of adequate vitamin D and calcium intake on musculoskeletal symptoms and long-term bone health is of particular relevance in postmenopausal breast cancer patients treated with aromatase inhibitors (Khan et al., 2010). These drugs are associated with increased risk for bone loss and fractures, in contrast to the bone protective effect of the selective oestrogen receptor (ER) modulator tamoxifen (Bauer et al., 2012). In this context exposure to sunlight (seasonal changes), population genetics, physical activity and nutritional factors are important considerations (Dawson-hughes & Harris, 2002; Grant & Boucher, 2017), particularly among postmenopausal breast cancer patients treated with aromatase inhibitors. These patients represent an important target group for the studies of genetic variation in the vitamin D pathway / receptor as the focus of this study.

The finding that genetic information may be insufficient to explain familial risk or predict treatment response, led to the development of a pathology-supported genetic testing (PSGT) platform for research translation (Kotze et al., 2015). This involves the generation of a genomics database using an integrated service and research approach. As explained on the informed consent form designed for this purpose, data can be extracted from the genomics database to 1) study the role of genetics in health and disease and to 2) provide information that could help clinicians improve the medical treatment of their patients. Comparative effectiveness studies of emerging genomic technologies performed together with standard pathology assessments may translate into immediate clinical benefits. This was clearly demonstrated by safe avoidance of chemotherapy in approximately 50% of early-stage South African breast cancer patients referred for the 70-gene MammaPrint test, based partly on a pre-screen step using immunohistochemistry (IHC) assessment of oestrogen receptor (ER),

progesterone receptor (PR) and human epidermal growth factor receptor-2 (HER2) status (Pohl et al., 2016). Application of PSGT furthermore facilitated interpretation of whole exome sequencing (WES) results in *BRCA1/2* mutation-negative breast cancer patients with tumour heterogeneity noted among affected members in the same family, including the index case referred for the MammaPrint test in 2008 (van der Merwe et al., 2017). Insight gained from these studies highlighted the importance of a clinical pipeline to define the target population most likely to benefit from aromatase inhibitor pharmacogenetic testing among South African breast cancer patients with known vitamin D levels (Baatjes et al., 2017).

The majority (55%) of postmenopausal hormone-positive South African patients treated with aromatase inhibitors in the study of Baatjes (2018) were found to be vitamin D deficient. This could have contributed to significant bone loss (>5%) linked to the *CYP19A1* rs10046 AA genotype implicated in a 12-month follow-up study. However, the cause for a similar degree of bone loss in a number of patients without this genotype, as well as osteopenia or osteoporosis detected in one-third of cases at baseline (Baatjes et al., 2019), remains to be identified in the same study cohort.

1.1 Rationale

Vitamin D deficiency correlates with poor outcomes in patients with the luminal A and luminal B breast cancer subtypes, while a similar association could not be demonstrated in patients with HER2-enriched or triple-negative breast cancer (Kim et al., 2011). Postmenopausal patients with the luminal-type breast cancer, which represents the most common form of this malignancy worldwide, were selected as the target population for the present study.

Conflicting results on the relationship between circulating vitamin D levels and breast cancer may be related to the inability to agree upon a generally accepted reference range for vitamin D concentration (Kennel et al., 2010; Holick 2009). This raised the possibility that detection of the genetic component of vitamin D status in breast cancer patients may improve risk stratification, as germline variants cannot be modified by the disease or vice versa. Different

approaches used to investigate the role of vitamin D in breast cancer include clinical trials, case-control and epidemiological studies in different geographical regions, and studies of mechanism to measure the range associated with high-risk clinical outcomes and cancer prevention.

1.2 Aim

The aim of the study was to determine the relationship between clinical characteristics, tumour histopathology and genetic variation underlying vitamin D metabolism in the study cohort. The objectives were:

1. To stratify the study population according to vitamin D levels documented in the research database.
2. Perform WES in DNA samples of vitamin D-deficient breast cancer patients with ultra-low levels (≤12 ng/mL) for comparison of their vitamin D-related genetic profile with vitamin D sufficient (>30 ng/mL) patients in the highest extreme of normal values.
3. Select clinically relevant single nucleotide variants (SNVs) based on the WES results and the literature study for genotype-phenotype association studies.
4. Perform variant classification and clinical interpretation for return of research results in eligible patients.

To meet these objectives, the research was conducted in four phases as shown in the research plan (**Figure 1.3).** Ultimately, new knowledge gained from this study will be incorporated into the PSGT platform in collaboration with the treating clinicians, to facilitate improved clinical management of hormone receptor-positive postmenopausal breast cancer patients treated with aromatase inhibitors.

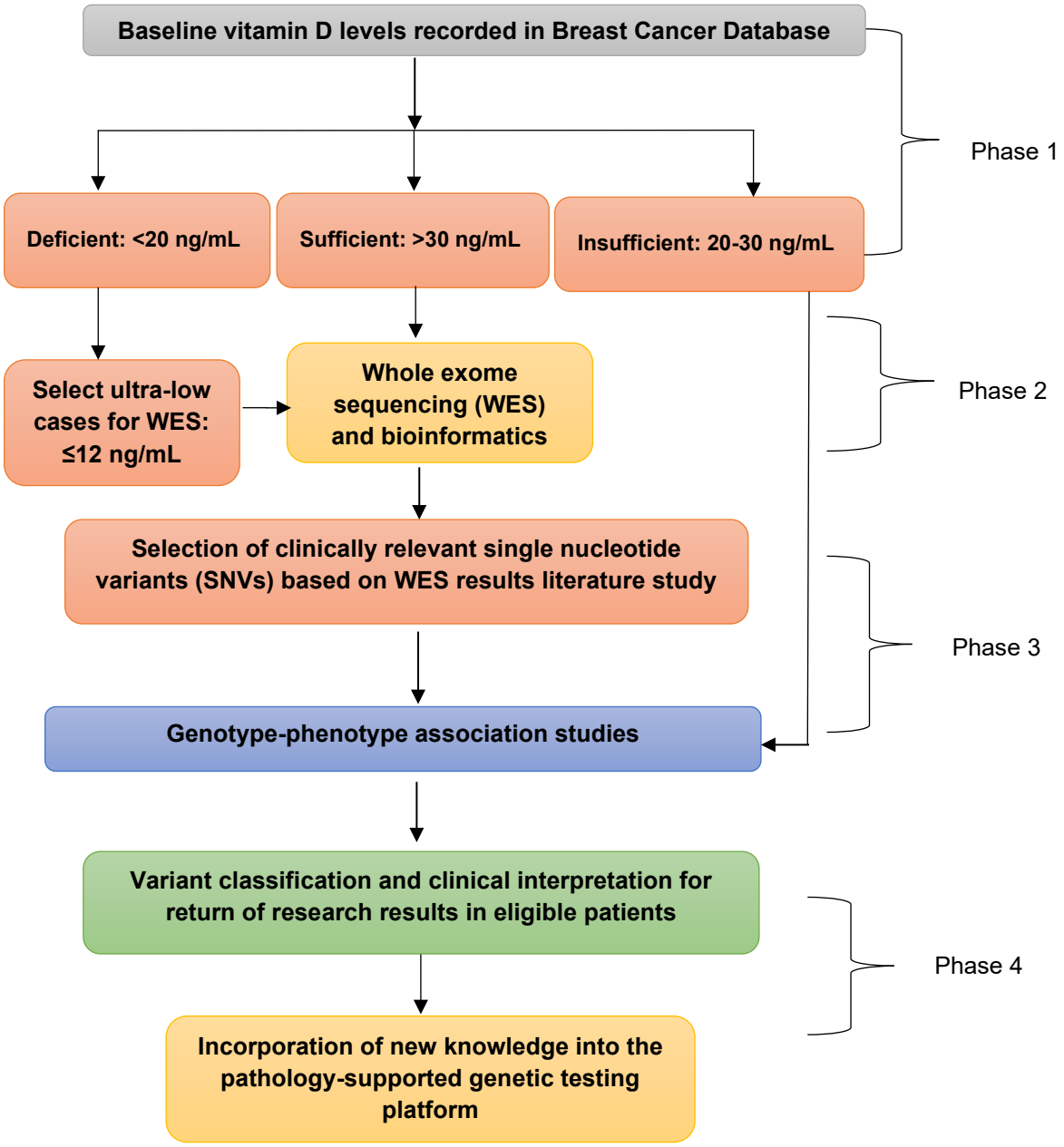


Figure 1.3: The research study plan. This research was conducted in four phases involving 1) extraction of relevant information from the genomics database, 2) whole exome sequencing (WES) at the Central Analytical Facility (CAF) of Stellenbosch University (Ion Proton), 3) extended genotyping of selected single nucleotide variants (SNVs) and long-range nanopore sequencing (MinION) at the Pathology Research Facility of Stellenbosch University, and 4) statistical analysis for interpretation of the data and incorporation of new knowledge into the pathology-supported genetic testing (PSGT) platform towards application of point-of-care technology.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Breast cancer is the leading cancer among women worldwide (Abulkhair et al., 2016). The age-standardized occurrence ratio per 100,000 ranges from 27 to 96 in Eastern Asia and Western Europe, respectively (Ferlay et al., 2015). The incidence is 26.8 per 100,000 in middle Africa, 30.4 in Eastern Africa, 38.6 in Western Africa, and 38.9 in Southern Africa. Genome-wide association studies (GWAS) have revealed distinct patterns in breast cancer patients of African ancestry compared to European populations, which underlie various biological pathways involved in the pathogenesis of breast cancer (Chen et al., 2013). Vitamin D levels are strongly influenced by environmental living conditions in various geographical locations, emphasising the importance of access to care and population differences in risk factors that might impact public health.

Sufficient circulating vitamin D levels activating the vitamin D receptor (VDR) may protect against cancer progression through regulation of cell division and apoptosis (Khammissa et al., 2018; Christakos et al., 2016). It was therefore suggested that optimal levels of vitamin D and pleiotropic effects of the *VDR* gene might be important in cancer prevention and survival outcome (Christakos et al., 2016). Strategies to optimize vitamin D levels are therefore important, especially in countries where cancer morbidity is increasing (Moukayed & Grant, 2013).

Epidemiological studies provide strong support for the role of vitamin D in increasing breast cancer survival outcome (Ordóñez-Mena et al. 2016;Garland et al. 2006; Tavera-Mendoza et al. 2017; Bandera Merchan et al. 2017). Vitamin D affects up to 5% of the human genome by regulating chromatin structure and gene expression. In a randomized, double-blind, single centre trial of vitamin D supplementation, Hossein-Nezhad et al. (2013) demonstrated at least

1.5-fold alteration in the expression of 291 genes with improved serum vitamin D concentrations. A significant difference in expression of 66 genes was noted between subjects with vitamin D levels below 20 ng/mL versus above 20 ng/mL at baseline, while gene expression was similar for both groups after two months of vitamin D supplementation. This study improved our understanding of the molecular fingerprints that explain non-skeletal benefits of vitamin D, which affects more than 160 genetic pathways linked to cancer, CVD and autoimmune disorders. Genes found to be important for DNA repair and transcriptional regulation include *TRIM27, CD83, COPB2, YRNA and CETN3* (Hossein-Nezhad et al. 2013).

Studies in Africa have predominantly concentrated on breast cancer as a single disease (Brinton et al., 2014), while next generation sequencing (NGS) defined at least 10 distinct molecular subtypes (Curtis et al., 2012). These include four major subtypes with different treatment requirements, initially discovered using microarray gene profiling (Perou et al., 2000): Luminal A, luminal B, HER2-enriched, and basal-type. Determination of oestrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) status are routinely performed using immunohistochemistry (IHC) as a proxy for these molecular subtypes in all newly diagnosed breast cancer patients. Familial breast cancer cases not explained by pathogenic mutations in the two major cancer causing genes, *BRCA1* and *BRCA2*, and differences in survival outcome, may be explained by low to moderate penetrance susceptibility alleles influenced by environmental factors (Gracia-Aznarez et al., 2013).

More than 80% of obese postmenopausal breast cancer patients in France were found to be vitamin D deficient (Bouvard et al., 2012). Similarly, only 7% of postmenopausal hormone receptor-positive South African patients treated at Tygerberg Hospital between 2014 and 2017 had sufficient vitamin D levels (Baatjes et al., 2019), in contrast to the majority of healthy individuals from the same population (Norval et al., 2016; Visser et al., 2019). Metabolic production and optimal vitamin D concentration in the body largely depend on the activity of the 1-alpha-hydroxylase enzyme *(CYP27B1)* and function of a group-specific component *(GC)*

protein, also known as the vitamin D-binding protein *(DBP)* (Thorne & Campbell 2008; Haussler et al. 2013). In a study performed by Batai et al. (2014), common vitamin D pathway gene variants revealed different effects on serum vitamin D levels in African Americans and European Americans. Replication of the phenotypic effect of six GWAS-identified single nucleotide variants (SNVs) confirmed the key role of *GC/DBP* in the vitamin D pathway across ethnic groups. Combining SNVs in the *GC/DBP*, *DHCR7/NADSYN1* and *CYP2R1* genes previously found to be associated with reduced vitamin D levels into a genotype risk score failed to show a significant effect on cancer risk in the Genome Health Study (Chandler et al., 2018). This may be due to the fact that breast cancer, which developed in 1560 individuals among 23 294 cancer-free cases followed up for 20 years, was studied as a single disease despite the presence of at least four major subtypes underpinned by genetic differences.

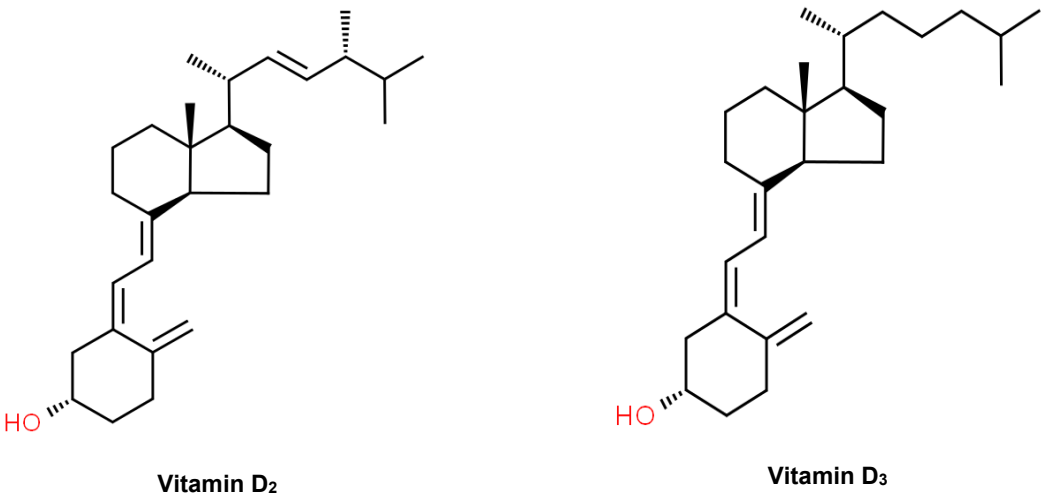
Conflicting findings on the relationship between vitamin D levels, genetic variation and clinical outcome can be ascribed to patient selection, different methods of estimating vitamin D levels, small sample size causing lack of statistical power, and limitations of single-gene genotyping methods applied. NGS is therefore used increasingly to simultaneously detect rare pathogenic mutations in high-risk genes and functional single nucleotide polymorphisms (SNPs) with minor allele frequency (MAF) greater than >1%. Detection of rare variants in moderate-risk genes with variable penetrance using NGS led to the term “single nucleotide variants (SNVs)” used in this study to describe gene variants of potential clinical significance across the spectrum of mutation penetrance or frequency. Interpretation challenges associated with variants of uncertain clinical significance (VUS) highlighted the need to take tumour pathology into account during variant classification. This led to development of a pathology-supported genetic testing (PSGT) algorithm, which confirmed the potential impact of genomic instability caused by dysfunction of the folate-homocysteine-methylation pathway (van der Merwe et al., 2017). A similar PSGT approach to study the role of vitamin D-related genes has not previously been performed in South Africa, after taking *BRCA1/2* mutation-status or the potential effect of other high-penetrance genes causing familial breast cancer into account.

Notably, germline pathogenic mutations in the cadherin-1 (*CDH1*) gene activated by VDR (Peñ et al., 2005), are associated with invasive lobular carcinoma (ILC), but not invasive ductal carcinoma of no special type (ICNST) (Lopes et al., 2012).

2.2 Structure and metabolism of vitamin D

2.2.1 Structure

Vitamin D is a steroid hormone that functions primarily in bone homeostasis (Atoum & Alzoughool, 2017). Levels of vitamin D are regulated by parathyroid hormones, and phosphate and calcium levels (Haarburger et al., 2009). Structural changes are considered to play an important role in different health conditions such as bone disease, cancer, coronary heart disease, dementia, infections, pregnancy, immune disorders and diabetes (Vimaleswaran et al., 2014; Ye et al., 2015; Pilz et al., 2013). Vitamin D exists in two forms namely vitamins D_2 (ergocholecalciferol) and D_3 (cholecalciferol) as shown in **Figure 2.1**



Vitamin D_2 **Vitamin D_3**

Figure 2.1: Structure of vitamin D2 and D3 forms (Source: chemspider.com)

The former is largely produced in plants while the latter is produced in mammals (Black et al., 2017). About 30% of vitamin D_2 can be obtained from dietary sources but few foods naturally

contain it (Palomer et al., 2008). The majority of vitamin D_3 is produced when human skin is irradiated by the ultraviolet component of sunlight and reacts with 7-dehydrocholesterol to form previtamin D_3, which is subsequently converted to vitamin D_3 (Wacker & Holick, 2013; Palomer et al., 2008). Both vitamin D_2 and D_3 will be referred to as vitamin D in the context of this thesis.

2.2.2 Metabolism

Both forms of vitamin D (D_2 and D_3) must be hydroxylated twice, first at position 25 in the liver and then at position 1 in the kidney, in order to yield the biologically active form, 1,25-dihydroxyvitamin D. Total 25 (OH) vitamin D, the main circulating form of vitamin D, has a half-life of approximately 2-3 weeks compared to 1,25 dihydroxyvitamin D that has a circulating half-life of 4-6 hours (Holick, 2009). Initiation of this cascade and binding of these metabolites to a carrier protein, the vitamin D binding protein (DBP), leads to the activation of the vitamin D receptor (VDR) in target organs such as skin, lungs, breast, intestine and prostate (Yousefzadeh et al., 2014; Bikle, 2014). Some cytochrome P450 mixed-function oxidases (CYPs) are important in vitamin D metabolism (Bikle, 2014), including cytochrome P450 family 2 subfamily R member 1 (*CYP2R1)*, cytochrome P450 family 24 sub family A member 1 *(CYP24A1)* and cytochrome P450 family 27 subfamily B member 1 *(CYP27B1)*.

2.3 Genes involved in vitamin D metabolism and signalling

2.3.1 The vitamin D binding protein

The DBP also known as a group-specific component (*GC*), is a polymorphic serum glycoprotein with several functions. In addition to the role in vitamin D transportation, DBP is an important regulator of vitamin D activity (Amadori et al., 2017). Based on the structure obtained from X-ray crystallography, DBP is composed of three domains: the first domain is made up of 10 α-helices, the second domain is similar to the first domain except for the replacement of a coil with helix 7, and the last domain is made up of only 4 helices (Rochel & Molnár, 2017).

Due to its relative stability, a higher percentage of circulating vitamin D are bound to these binding proteins (Yousefzadeh et al., 2014). However, some physiological factors such as: race, genetic polymorphisms, age and gender, obesity, pregnancy and oestrogen as well as pathologies such as liver disease, renal disease, diabetes, primary hyperparathyroidism, cancer, human immunodeficiency virus (HIV) and acute inflammation can affect DBP and binding affinity of the circulating vitamin D. *DBP/GC* gene variants including *rs4588, rs7041, rs16846876, rs2282679, rs12512631, rs17467825*, and *rs842999* have been found to be associated with reduced vitamin D serum levels. Studies in different populations have confirmed the gene effect across ethnic groups (Nissen et al., 2014).

2.3.2 Cytochrome P450 family 2 subfamily R member 1

Cytochrome P450 family 2 subfamily R member 1 *(CYP2R1)* is responsible for the first hydroxylation in the vitamin D activation pathway. Genetic variation in the *CYP2R1* gene may result in vitamin D deficiency. From different population studies, *CYP2R1* variants *(rs1562902, rs7116978, rs12794714, rs10741657,* and *rs10766197)* were found to be positively associated with serum vitamin D concentrations (Nissen et al., 2014).

2.3.3 Cytochrome P450 family 24 subfamily A member 1

Cytochrome P450 family 24 subfamily A member 1 *(*CYP24A1*)* and family 27 subfamily B member 1 *(*CYP27B1*)* are the two major enzymes involved in vitamin D metabolism (Srilatha Swami, Krishnan, & Feldman, 2011). CYP24A1 catalyzes the conversion of 25-hydroxyvitamin D and 1,25-dihydroxyvitamin D into 24-hydroxyvitamin D, the degraded vitamin D molecule in the target tissues (Jones et al., 2012; Swami et al., 2011). It regulates vitamin D concentration and clearance of its inactive form (24-hydroxyvitamin) (Jones et al., 2012; Nissen et al., 2014). *CYP24A1* variants (*rs6013897* and *rs17217119*) have been suggested to have an association with vitamin D levels (Coussens et al., 2015).

2.3.4 Cytochrome P450 family 27 subfamily B member 1

CYP27B1 is involved in the hydroxylation of the second step in vitamin D activation pathway. *CYP27B1* variants *rs10877012, rs4646536* and *rs703842* are associated with vitamin D levels in some population groups or disease conditions, but not in others. For example, the T allele of variant *rs10877012* was shown to be associated with low serum vitamin D levels in African Americans, but not in Caucasians (Signorello et al., 2011). The C allele of *rs703842* and T alleles of *rs10877012* and *rs4646536* found to be in linkage disequilibrium, were associated with decreased vitamin D levels in Canadian multiple sclerosis patients and the Han Chinese population (Jiang et al., 2016; Zhuang et al., 2015).

2.3.5 The vitamin D receptor

Vitamin D receptor (VDR) protein is a member of the nuclear receptor superfamily, which plays a significant part in the biological actions of vitamin D (Wang et al., 2012). It functions primarily as a transcriptional activator of many genes that are expressed in most human tissues, normal breast and breast cancer inclusive (Haussler et al. 2013; Krishnan & Feldman, 2011). VDR undergoes a conformational change when vitamin D binds to it that enables it to bind to retinoid X receptor (RXR) and form heterodimer that reacts with vitamin D-responsive elements in the promoter region of target genes and modify their expression (Mackawy et al. 2014). Strauss (2014) postulated a membrane VDR (mVDR) existence and suggested its receptiveness for vitamin D effects that do not involve gene expression such as activation of protein kinase as well as a rise in intracellular calcium and cGMP levels (Strauss, 2014). The most frequently studied variants of VDR are located at the 3' end of the gene, including *rs7975232 (ApaI), rs1544410 (BsmI), rs2228570 (FokI), and rs731236 (TaqI)* (Uitterlinden et al., 2004). The allele frequencies of these and other variants in the *VDR* gene may differ between populations and their effects on the vitamin D endocrine system, gene regulation or protein structure remains uncertain in relation to complex conditions and traits, such as osteoporosis and vitamin D deficiency (Uitterlinden et al., 2004). VDR has also been implicated in the regulatory function of vitamin D, which requires careful consideration of both the genotype and

epigenome to determine the root cause of population risk (Andraos et al., 2011), which proved to be an imperfect proxy for targeting multiple environmental and genetic factors involved in disease development and progression or response to treatment (Collins, 2004). **Table 2.1** provides a summary of commonly studied variants in vitamin D-related genes.

Table 2.1: Vitamin D-related genes and variants previously studied in relation to functional effects on gene regulation and protein function

Vitamin D Genes and Variants	Location	Amino acid change	Role in vitamin D metabolism	References
CYP2R1 *rs10741657*	Promoter region	N/A	Hepatic hydroxylation	(Coussens et al., 2015)
CYP24A1 *rs6013897* *rs17217119* *CYP27B1* *rs12368653* *rs703842*	3' downstream	N/A	Catabolism	(Coussens et al., 2015) (Zhuang et al., 2015)
DBP *rs4588* *rs7041* *VDR* *rs2228570 (FokI)* *rs1544410 (BsmI)* *rs7975232 (ApaI)* *rs731236 (TaqI)*	 Exon 11 Exon 12 Exon 2 Intron 8 Intron 8 Exon 9	 T455K D451E M1K N/A N/A I402I	Transportation	(Coussens et al., 2015) (Lee et al., 2016) (Lombard et al., 2006)

CYP2R1: cytochrome P450 family 2 subfamily R member 1; CYP24A1: cytochrome P450 family 24 subfamily A member 1; DBP: vitamin D binding protein; VDR: vitamin D receptor; CYP27B1: cytochrome P450 family 27 subfamily B member 1; N/A: not applicable

2.4 Vitamin D and epigenetics

Epigenetics is defined as a process that modifies gene action without altering the DNA sequence (Fetahu et al., 2014). Vitamin D activation of VDR in metabolic activities controls epigenetic mechanisms (Karlic & Varga, 2011) to the extent that dysregulation can lead to the development of cancer and related conditions. Epigenetic processes involved in the regulation of gene expression includes DNA methylation and covalent modification of histones by methylation, with DNA methylation identified as a major epigenetic factor influencing gene expression (Moore et al., 2013).

The enzymes DNA methyl transferases and ten-eleven translocation proteins enable changes in DNA methylation, while acetyltransferases, deacetylases, methyl transferases, and demethylases regulate covalent histone modifications (Moore et al., 2013). DNA methylation only occurs at the cytosine residue-and can change the functional state of regulatory regions, by not base pairing of cytosine conferring epigenetic marks on specific sites (Schübeler, 2015). It is involved in different forms of stable epigenetic repression include X chromosome inactivation, imprinting and silencing of repetitive DNA (Jones, 2012). The discovery of the VDR highlighted the important role of vitamin D in gene regulation (Fetahu et al., 2014). Therefore, vitamin D may play a significant role in regulating epigenetic events in preventing and / or treating tumorigenesis and chronic diseases (Fetahu et al., 2014).

2.5 Laboratory testing of vitamin D

The standard method for vitamin D analysis is liquid chromatography-tandem mass spectrometry (LC-MS/MS), however most laboratories use immunoassay (Avenell et al., 2019). The difficulty with harmonizing LC-MS/MS and immunoassay is a major obstacle affecting the standardization of vitamin D determination (Fraser & Milan, 2013). Also, there is no agreement yet regards to the optimal serum vitamin D levels for better health and the acceptable level of deficiency is also being debated (Anandabaskar et al., 2018). Based on its measurement for healthy bone maintenance, the United States Endocrine Society has classified vitamin D levels into 3 categories using the following cut-offs: deficient: <20 ng/mL, insufficient: 21-29 ng/mL and sufficient: ≥30 ng/mL (Holick, 2009). However, the cut-off points of the Institute of Medicine classified vitamin D levels are deficient <12 ng/mL, insufficient 12-20 ng/mL and sufficient >20 ng/mL. However, we adopted the United States Endocrine Society classification system for this study because it is universally accepted. A detailed population study of the genetics of vitamin D levels in association with specific disease may help us understand the importance of vitamin D.

2.6 Factors affecting vitamin D levels

Vitamin D levels are affected by age (Atoum & Alzoughool, 2017), gender, pigment genes (Datta et al., 2019), latitude and environmental factors such as seasonal changes, BMI, physical activity, smoking and alcohol consumption (Shi et al., 2014). Engelman et al (2008) reported increased vitamin D serum levels in Hispanics from San Luis Valley and decreased serum levels in Hispanics from San Antonio and in African Americans from Los Angeles (Engelman et al., 2008). The observed higher vitamin D levels in the Hispanics in San Luis Valley when compared with those in the San Antonio were ascribed to more European ancestry and less African ancestry in this Hispanic population (Engelman et al., 2008). Yao et al (2017) also compared serum levels of vitamin D and DBP among a cohort of African Americans and European Americans, to identify determinants of vitamin D concentrations (Yao et al., 2017). They reported that African American women have lower serum vitamin D levels than European American women, however, both groups have similar DBP concentrations (Yao et al., 2017).

The skin pigment in humans is determined by the amount, type, and distribution of melanin, which varies in human populations and is of importance in vitamin D synthesis (Parra, 2007). Low vitamin D levels observed among darker-skinned populations are related to variation in pigment genes (Datta et al., 2019). With careful data interpretation, pigment genes may explain the diversity of many human traits (Parra, 2007). **Table 2.2** shows a list of pigment genes that have previously been studied. Encouragement to shift the focus from classification of populations based on ethnicity and race to inherent genetic individuality for effective optimization of medication in individual patients has led to the suggestion of developing a panel of “ancestry-informative” markers that classify and provide detailed information about the geographic region and its influence on an individual’s ancestors (Zhang & Finkelstein, 2019)

Table 2.2: Summary of pigment genes and common genetic variants associated with vitamin D levels

Research article title / Reference	**Pigment genes/variants**	**Case/control**	**Country**
Pigment genes not skin pigmentation affect UVB-induced vitamin D (Datta et al., 2019)	*ASIP (rs4911414, rs1015362), SCL24A4 (rs128963990), MTAP, MIR196A29, SCL45A2 (rs28777, rs16891982)*	C (40)	Denmark
Human pigmentation, cutaneous vitamin D synthesis and evolution: Variants of Genes Involved in Skin Pigmentation Are Associated with 25(OH)D Serum Concentration (Rossberg et al., 2016)	*ATP7A, DTNBP1, BLOC1S5, PLDN, PMEL, RAB27A, MYO5A, MLPH, MC1R (rs1805007, rs1805008), MITF, PAX3, SOX10, DKK1, RACK1, CNR1*	Cohort study (2974)	Germany
A Closer look at evolution: Variants of Genes Involved in Skin Pigmentation, Including *EXOC2, TYR, TYRP1, and DCT*, Are Associated with Vitamin D Serum Concentration (Saternus et al., 2015)	*EXOC2, TYR (rs1126809, rs1042602, PRKACG, EDN1, TYRP1 (rs1408799), MITF*	Cohort study (2970)	Germany
Genetic Ancestry, Skin Reflectance and Pigmentation Genotypes in Association with Serum Vitamin D Metabolite Balance (Wilson et al., 2011)	*SLC45A2 and SLC24A5*	Case/control (50/50)	United States of America
A polymorphism in IRF4 affects human pigmentation through a tyrosinase-dependent MITF/TFAP2A pathway (Praetorius et al., 2013)	*IRF4 (rs12203592)*	Not provided	Not provided
A global view of *the OCA2-HERC2* region and pigmentation (Donnelly et al., 2012)	*OCA2 (rs7495174)*	Case (3432)	Global

Sufficient circulating vitamin D levels may protect against cancer through the regulation of cell division and apoptosis (Fourie et al., 2018). One possible reason for inconsistent study results may include different assays used for serum vitamin D estimation. Environmental factors such as (i) dietary pattern of foods containing vitamin D and supplementation and (ii) varying

degrees of exposure to sunlight due to seasonal changes that differ among populations are also important contributing factors (Spiro & Buttriss, 2014). The use of Mendelian randomization studies to determine the role of genetic variation was inconsistent. Little evidence was provided that a vitamin D multi-polymorphism score increased the risk of seven forms of cancer, with an odds ratio of 0.89 (95% confidence interval 0.77 to1.02) for breast cancer per 25 nmol/L increase in genetically determined vitamin D levels (Dimitrakopoulou et al., 2017). Another study reported no evidence to support an association between vitamin D and risk of breast cancer with an odds ratio of 1.02 (95% confidence interval: 0.97-1.08) per 25 nmol/L increase (Jiang et al., 2018).

2.8 Cancer and global trends

Vitamin D levels may influence cancer development, progression, survival outcome and drug response. Therefore, an in-depth understanding of all these aspects is important. Cancer can be defined as a disease of abnormal cell growth when compared with the usual normal cell division (Huang, et al., 2018). The most common global causes for cancer-related death are lung, colorectal, stomach, liver and breast cancers (Ferlay et al., 2018). It is now a common disease in both developed and developing countries. Aging, obesity, physical inactivity, smoking, use of hormone replacement therapy (HRT), infection, radiation and chemical exposure, increase the incidence of cancer globally (Golemis et al., 2018).

2.8.1 Breast cancer

Breast cancer poses a great challenge to health, being the leading cancer among women worldwide (Bray et al., 2018). It has been reported to be on the increase in the Sub-Saharan Africa including Nigeria and South Africa, with survival rates much poorer than those in developed countries (Farmer et al., 2010). Breast cancer is a heterogeneous disease which varies significantly among different patients (inter-tumour heterogeneity) as well as within each

individual's tumour (intratumour heterogeneity) (Turashvili & Brogi, 2017). However, initial studies in Africa concentrated on breast cancer as a single disease (Brinton et al., 2014). Treatment as a single disease and presentation with the advanced stage of the disease makes survival very low on this continent (Ogunkorode et al., 2017). The primary factors responsible for the late presentations are lack of awareness, misconceptions about breast cancer causes, and treatment outcomes (Ogunkorode et al., 2017). Younger age at presentation is also thought to contribute to higher mortality rates (Newman, 2015).

2.8.2 Breast cancer risk factors

Generally, the risk factors that are associated with breast cancer are: gender, aging, family history, gene mutations, reproductive factors, oestrogen exposure and lifestyle factors (Feng et al., 2018; Sun et al., 2017; Kamińska et al., 2015). Being female is a key risk factor for breast cancer (Feng et al., 2018) with a lower occurrence in males (Greif et al., 2012; Miao et al., 2011). Early menarche, late menopause, late age-at-first pregnancy and low parity are factors that can increase the breast cancer risk (Sun et al., 2017). Studies have reported increased breast cancer risk by up to 3% in delayed menopause (Surakasula et al., 2014; Matthew Stenger, 2013), while additional births or delayed menarche decrease breast cancer risk (Sun et al., 2017).

Oestrogen, either endogenous or exogenous, is an important risk factor for breast cancer development (Sun et al., 2017). Increased endogenous oestrogen levels have been strongly associated with increased risk of breast cancer in postmenopausal women (Key et al., 2002), whereas it is associated with reduced breast cancer risk in premenopausal women (Health, 2013). The source of exogenous oestrogen is mainly through HRT, usually for postmenopausal women (Sun et al., 2017). The use of HRT has been shown to increase breast cancer risk as reported by The Million Women study in the United Kingdom (UK) (Beral & Million Women Study Collaborators, 2003) and in another study from Asia (Liu et al., 2016).

There has been a growing number of reports that regular physical activity in postmenopausal women may reduce the risk of breast cancer (Niehoff et al., 2019; Feng et al., 2018). Possible reasons could be because activity levels affect body weight, inflammation, hormones, and energy balance (Feng et al., 2018). Lifestyle factors such as excessive alcohol (Jung et al., 2016), significant overweight or obesity (Feng et al., 2018), and smoking (Makarem et al., 2013) also increase the risk of breast cancer. However, there are controversies regarding dietary fat intake (Makarem et al., 2013; Kotepui, 2016). Khodarahmi and Azadbakht (2014) based their study on the categorization of fats into saturated and unsaturated, suggested that consumption of unsaturated fatty acids and reduction of saturated fatty acids may be beneficial to reduce breast cancer risk (Khodarahmi & Azadbakht, 2014). Although fat tissues produce some small additional amount of oestrogen, the ovaries produce amounts that are sufficient during the premenopausal phase (Sun et al., 2017; Kamińska et al., 2015). At the postmenopausal phase, the ovaries cease to produce oestrogen so fat tissues take over its production (Feng et al., 2018). Being overweight or obese is also associated with increased circulating insulin levels which are linked to cancers, breast cancer inclusive (Feng et al., 2018). **Figure 2** shows a simplified schematic presentation of the role of obesity in breast cancer development.

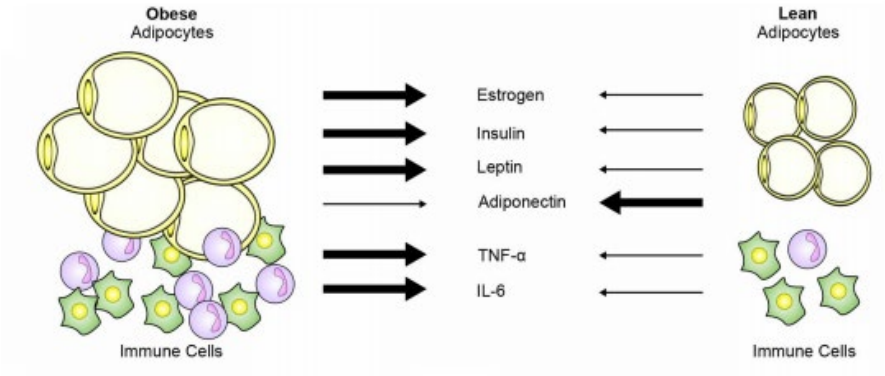


Figure 2.2: Cancer related molecular mechanisms driven by obesity are associated with elevated aromatase expression mediated partly by increased levels of prostaglandin and TNF-α. Adapted from: Molecular mechanisms of the preventable causes of cancer in the United States (Adapted with permission from Golemis et al., 2018).

Family history is also a significant risk factor for breast cancer (Brewer et al., 2017; Sun et al., 2017). A UK study including more than 100,000 participants reported that women with one first-degree relative with breast cancer had a 1.75-fold increased risk of developing this disease. Furthermore, the risk increased 2.5-fold in women with two or more first-degree relatives with breast cancer (Brewer et al., 2017).

Generally, about 25% of breast cancer cases are hereditary (Balmaña et al., 2011). *BRCA1* and *BRCA2* alongside *TP53, PTEN, STK11* and *CDH1* being classified as high penetrance genes (Han et al., 2017; Shiovitz & Korde, 2015). Patients with pathogenic mutations in these gene have an 40-80% lifetime risk of breast cancer (Han et al., 2017). Also, moderate-penetrance mutations in the *CHEK2, ATM, BRIP1, PALB2, RAD51C,* and *RAD50* genes have been identified as a cause or contributing factor for breast cancer development (Han et al., 2017; Shiovitz & Korde, 2015). **Table 2.3** shows the summary of well-established genes associated with breast cancer.

Table 2.3: Summary of genetic abnormalities previously associated with familial breast cancer

Gene	Location	Genetic risk	Abnormality in breast cancer	References
BRCA1	17q21.31	Increased breast cancer risk	Cell cycle checkpoint dysregulation, abnormal duplication of centrosome, genetic irregularity and finally cell death	(Sun et al., 2017) (Dine & Deng, 2013)
BRCA2	13q13.1	Increased breast cancer risk	Invasive ductal carcinomas of no special type, showing a luminal phenotype	(Sun et al., 2017) (Bane et al., 2007)
TP53	17p13.1	Increased risk of several forms of cancers, including breast cancer	Li-Fraumeni syndrome and subsequently an increased susceptibility in developing breast cancer in about 30% of carriers	(Frey et al., 2017) (Damineni et al., 2014)
PTEN	10q23.31	Risk of developing breast cancer	Cowden syndrome, presenting with hamartomas and benign tumours, macrocephaly, high-flow vascular malformations, and plantar keratosis	(Feng et al., 2018) (Sun et al., 2017) (Frey et al., 2017)

Gene	Location	Genetic risk	Abnormality in breast cancer	References
STK11	19p13.3	Increased risks of certain types of cancers, including breast cancer	Peutz-Jeghers syndrome presenting with small bowel hamartomas and skin discoloration in buccal mucosa, lips, fingers and toes	(Feng et al., 2018) (Frey et al., 2017)
CDH1	16q22.1	Hereditary diffuse gastric cancer with an increased risk of invasive lobular breast carcinoma	Specifically associated with invasive lobular carcinoma	(Feng et al., 2018) (Sun et al., 2017)
CHEK2	22q12.1	Increased breast cancer risk	Inability to repair DNA or undergo cell cycle arrest in response to damage	(Feng et al., 2018) (Frey et al., 2017)
ATM	11q22.3	Increased breast cancer risk	Inability to suspend cell division and initiate DNA repair	(Sun et al., 2017) (Frey et al., 2017)
BRIP1	17q23.2	Increased breast cancer risk	Inability to repair DNA	(Frey et al., 2017)
PALB2	16p12.2	Increased breast cancer risk	Fanconi's anaemia, presenting with bone marrow failure, developmental anomalies, and hereditary breast cancer development	(Feng et al., 2018) (Frey et al., 2017)
RAD50	5q31.1	Increased breast cancer risk	Inability to respond to double strand breaks in DNA and to initiate DNA repair mechanisms	(Frey et al., 2017)
RAD51C	17q22	Increased breast cancer risk	Inability to produce proteins involved in DNA recombination repair	(Frey et al., 2017)

The development of advanced laboratory techniques has led to improved sequencing capabilities to identify moderate to low penetrance genes with genetic variants that may contribute to breast cancer risk, prognosis and response to treatment in a polygenic manner (Shiovitz & Korde, 2015). Common genetic variants are often used to evaluate genetic variability (Shiovitz & Korde, 2015) and may be population-specific, implying a role in a large proportion of that population as a whole, compared to the high impact of high-penetrance mutations in a relatively small number of patients at the family / individual level (Weber & Nathanson, 2000). Among such low penetrance genes are several members of the

cytochrome P450 superfamily; *CYP1A1, CYP17, GSTs* (*Glutathione S-transferases*) and *NAT2* (*N-acetyl transferase 2*) (Sillanpaa, 2007; Weber & Nathanson, 2000). Low penetrance genes associated with vitamin D metabolism and signalling include *VDR, CYP27B1, CYP24A1, GC* and *DHCR7* (Lopes et al., 2010; Nissen et al., 2014).

2.8.3 Breast cancer classification

Breast cancer can be classified into three different types based on pathology, invasiveness and prevalence (Feng et al., 2018). On a pathological basis, breast cancer can start from different areas of the breast like the ducts, the lobules, or the tissue in between. Hence, it can either be carcinomas or sarcomas (Feng et al., 2018). The former is more frequent and based on its invasiveness, and can be grouped into non-invasive (or in situ) and invasive (Malhotra et al., 2010). Based on growth patterns and cytological features, non-invasive carcinoma has been further categorized as either ductal or lobular carcinoma (Malhotra et al., 2010).

2.8.3.1 Non-invasive carcinomas

Ductal carcinoma in situ (DCIS) is expressed as an abnormal growth in the breast epithelial cells but is still intact in the ducts and lobules (Allred, 2010). DCIS is more frequent than LCIS and can be categorized into 5 distinct subtypes based on the features expressed by the tumour, namely comedo, cribiform, micropapillary, papillary and solid. DCIS tumours express mixed features of these subtypes (Makki, 2015). According to their morphology, DCIS is generally divided into low-grade (Lakhani et al., 2012), intermediate-grade (Makki, 2015; Tavassoli & Devilee, 2003) and high-grade (Makki, 2015; Tavassoli Fattaneh & Devilee, 2003). Limitation of this classification is its inability to utilize/incorporate confirmed molecular markers for prognostic importance (Malhotra et al., 2010; Allred, 2010).

Lobular carcinoma in situ (LCIS) is the multiplication of small, fairly uniform and loosely cohesive cells within the breast lobules with or without involvement of terminal ducts (Makki, 2015). Over a period of time, women with LCIS have a high risk of developing invasive carcinoma (Tavassoli & Devilee, 2003). It has no specific distinguishing features, and is

commonly found incidentally in breast specimens taken for other reasons. About 70% of cases are multicentric while approximately 30%–40% are bilateral (Rosai & Ackerman, 2011). It is usually presented as round, small-to-medium sized cells with normochromatic nuclei that filled lobules in a non-cohesive pattern, changes such as pleomorphism, mitosis, and necrosis are absent, intracellular mucin droplets are seen occasionally having signet ring nuclei (Makki, 2015; Hanby & Hughes, 2008). While LCIS lacks immunohistochemical markers; E-cadherin, β-catenin expressions and the positivity for high molecular weight (HMW) keratin (Mastracci et al., 2005), DCIS is positive for E-cadherin and β-catenin and expresses a low or no high molecular weight (HMW) keratin (Makki, 2015).

2.8.3.2 Invasive carcinomas

Based on histological features, carcinomas have been classified as invasive/infiltrating ductal, invasive lobular, ductal/lobular, mucinous (colloid), tubular, medullary and papillary carcinomas (Malhotra et al., 2010). Invasive ductal carcinoma (IDC) is the most common subtype accounting for 70–80% of all invasive lesions (Li et al., 2005). About 75% of IDC lack morphological features to be classified as specific histological types and are referred to as otherwise specified. The term “no special type” (NST) has been internationally accepted to differentiate it from specific-type tumours (Lakhani et al., 2012).

Invasive lobular carcinoma (ILC) is the second most common histologic subtype, it accounts for about 10% of all breast cancers (Barroso-Sousa & Metzger-Filho, 2016); (Li et al., 2003). It is different from the most common subtype, IDC with respect to epidemiology, molecular alterations, clinicopathology and natural history (Rakha & Ellis., 2010). The occurrence of ILC is on the increase, mostly in postmenopausal women, which may likely be associated with HRT (Makki, 2015; Tavassoli & Devilee, 2003). The cause of non-cohesive characteristics in ILCs is due to lack of E-cadherin expression which is responsible for cell–cell adhesion properties (Barroso-Sousa & Metzger-Filho, 2016). About 90% of ILCs lack E-cadherin protein expression which is the main distinguishing characteristic of ILCs (Barroso-Sousa & Metzger-

Filho, 2016; Makki, 2015). E-cadherin deregulation is driven by genomic alterations (Rosai & Ackerman, 2011) that target the *CDH1* gene (Barroso-Sousa & Metzger-Filho, 2016); Makki, 2015; McCart Reed et al., 2015).

Unlike non-invasive carcinomas, molecular biomarkers ER, PR and HER2 are incorporated in classification of invasive carcinomas (Malhotra et al., 2010). The use of these biomarkers demonstrates their importance in making relevant important clinical decisions (Turashvili & Brogi, 2017; Karen et al., 2010). The status of these markers is beneficial in determining which patients are most likely to respond to targeted therapies such as tamoxifen or aromatase inhibitors for hormone receptor-positive and trastuzumab or lapatinib for HER2-positive breast cancer (Makki, 2015; Rakha et al., 2010).

2.8.4 Molecular classification of breast cancer subtypes

Due to breast cancer heterogeneity researchers have studied the molecular structure for a better understanding to improve treatment strategies (Feng et al., 2018; Makki, 2015). Perou and colleagues (2000), on the basis of similarities in gene expression profiles with the aid of microarray technology, classified breast cancer into four major subtypes as luminal A, luminal B, HER2-enriched, and basal-like (Perou et al., 2000). The fifth “normal breast–like” category has not been reproducibly defined (Allison, 2012). The four major subtypes (**Table 2.4**), initially discovered by gene-expression profiling, were confirmed to be valid entities when other platforms such as genomic DNA copy number arrays, DNA methylation, exome sequencing, microRNA sequencing, and reverse phase protein assays were used in comparative studies (The Cancer Genome Atlas Network, 2012).

Table 2.4: Tumour characteristics of breast cancer subtypes and treatment options

Subtypes	Prevalence	Descriptions	Treatment and Prognosis
Luminal A	40%	ER+ and/or PR+, HER2- and low Ki-67	Good prognosis, typically treated with hormonal therapy
Luminal B	<20%	ER+ and/or PR+, HER2+ or HER2-, high Ki-67	Slightly worse prognosis, requires addition of chemotherapy to endocrine treatment
HER2-enriched	10-15%	ER-, PR-, and HER2+	Poor prognosis, successfully treated with targeted therapies including Herceptin (trastuzumab), Perjeta (pertuzumab), Tykerb (lapatinib), and Kadcyla (T DM1 or ado-trastuzumab emtansine).
Basal-like (TNBC)	~20%	ER-, PR-, and HER2-	Chemotherapy, PARP inhibitors (e.g. olaparib)
Normal-like	~5%	ER+ and/or PR+, HER2- and low Ki-67	Prognosis similar to luminal A subtype

It is important to know that the HER2-enriched molecular subtype is not synonymous with clinically HER2-positive breast cancer. While about 50% of clinical HER2-positive breast cancers are HER2-enriched, the remaining 50% can include any molecular subtype, but are mostly HER2-positive luminal A or B subtypes. About 30% of HER2-enriched tumours are clinically HER2 negative. The basal subtype overlaps largely with triple negative breast cancer (TNBC) found to be more common in younger women (<40 years) with *BRCA1* gene mutations (Feng et al., 2018).

2.9 Vitamin D and breast cancer risk

Studies on the risk implications of vitamin D deficiency should be based on the understanding that breast cancer is a heterogeneous disease consisting of different subtypes requiring specific treatment strategies. A recent systematic review on the impact of serum vitamin D levels on receptor status identified an increased risk for TNBC, hence routine screening should be implemented for early detection of vitamin D deficiency focused on further studies to confirm if the association is causative (Tommie et al., 2018). A hospital-based case–control study furthermore reported a reduced risk of breast cancer in premenopausal women with

higher vitamin D levels, which proposes that vitamin D supplementation possibly will be beneficial to reduce breast cancer risk (Yao et al., 2011).

2.10 Vitamin D and breast cancer subtypes

Low vitamin D levels are common in postmenopausal breast cancer patients at diagnosis (Machado et al., 2019; Hatse et al., 2012), with increased breast cancer survival rates associated with vitamin D supplementation (Madden et al., 2018). A Korean study reported that vitamin D deficiency correlated with poor outcomes in patients with luminal A and luminal B breast cancer (Kim et al. 2011). Conversely this association was not found in patients with TNBC or the HER2-enriched subtype. Park and colleagues (2015) reported the strongest association between TNBC and low serum vitamin D levels (Park et al., 2015), which agrees with previous studies (Abbas et al 2009; Bertone-Johnson et al 2005). Reports from South Africa and Namibia revealed a relatively large percentage of women with ER-positive tumours (Dickens et al., 2014). Studies performed in West and Central Africa revealed that breast cancer occurs at younger ages and usually presents with aggressive features, such as high-grade, advanced stage and triple-negative phenotype (negative for ER, PR and HER2) (Brinton et al., 2014). However, in Southern and Eastern African countries, hormone receptor-negative tumours account for a lower proportion of all breast cancers (Hadgu et al., 2018; Joffe et al., 2018). In Ethiopia, luminal A breast cancer was found to be the most common subtype based on the following distribution: 40% luminal A, 26% luminal B, 10% HER2-enriched and 23% TNBC (Hadgu et al., 2018). Differences in the distribution of molecular subtypes of breast cancer in Africa were attributed to geographic factors, which among other factors have led to difficulty in treatment of breast cancer on the African continent.

With these varying outcomes, a better understanding of the clinical relevance of variants identified by NGS, ranging from low- to moderate- and high-penetrance mutations, is of utmost importance to identify patients at increased risk of poor clinical outcome.

2.11 Vitamin D and breast cancer treatment

Compared to BRCA2, BRCA1-related tumours predominantly associated with TNBC exhibit lower levels of nuclear VDR. At present aromatase inhibitors are the most effective treatment for endocrine-responsive breast cancers in postmenopausal women. Anastrozole, letrozole and exemestane have greater effectiveness in treating hormone-responsive postmenopausal breast cancers (Dixon, 2014). Randomized clinical trials also corroborate that these aromatase inhibitors are more effective than tamoxifen (Haque et al., 2012) and may possibly be effective in tamoxifen-resistant advanced breast cancer. Nevertheless, breast cancer treatment with aromatase inhibitors accelerates bone loss by reducing circulating levels of oestrogen (Ingle et al., 2010). Aromatase inhibitors may induce hypoestrogenaemia, thereby increasing the potential for osteoporosis in the postmenopausal population already at increased risk due to advanced age. Women with breast cancer may be at greater risk of osteoporosis, a phenomenon that may be related to genetic factors common to both diseases. The adverse side-effects of aromatase inhibitors may affect compliance to treatment and thereby reduce its efficacy (Reid et al., 2008; Aapro & Coleman, 2012).

To reduce the side effects and increase treatment efficacy, Prieto-Alhambra et al (2011) proposed that vitamin D supplementation can protect against the effect of bone loss among patients on aromatase inhibitors (Prieto-Alhambra et al., 2011). Although outcomes from randomized controlled trials of vitamin D supplementation have been inconclusive for various diseases, some studies reported an advantage of supplementation (Sperati et al., 2013; Redaniel et al., 2014). Other randomized clinical studies did not observe any statistically significant difference (Barbarawi et al., 2019; Manson et al., 2019; Pittas et al., 2019). Vitamin D regulates aromatase secretion in bone and enables increased production of aromatase mRNA (Yanase et al., 2003), as supported by *in vitro* and *in vivo* studies (Krishnan et al., 2010; Swami et al., 2011). The finding that reduced levels of vitamin D may be more tolerable in African Americans than in individuals of European descent in relation to bone health (Mckee

et al., 2018), may be due to differential effects on VDR expression influenced by both environmental and genetic risk factors (Signorello et al., 2011).

Calcium and vitamin D supplementation form an important aspect of the clinical pipeline developed by Baatjes et al. (2017) for clinical management of hormone receptor-positive postmenopausal South African breast cancer patients at increased risk for osteoporosis. The PSGT approach applied in this pharmacogenetics study incorporating advanced WES for selection of clinically relevant SNVs, supported the research translation process shown in **Figure 2.3**. Further studies are warranted to identify causal factors of vitamin D deficiency that may be targeted towards prevention of osteoporosis as a major co-morbidity of aromatase inhibitor treatment (Baatjes et al., 2019).

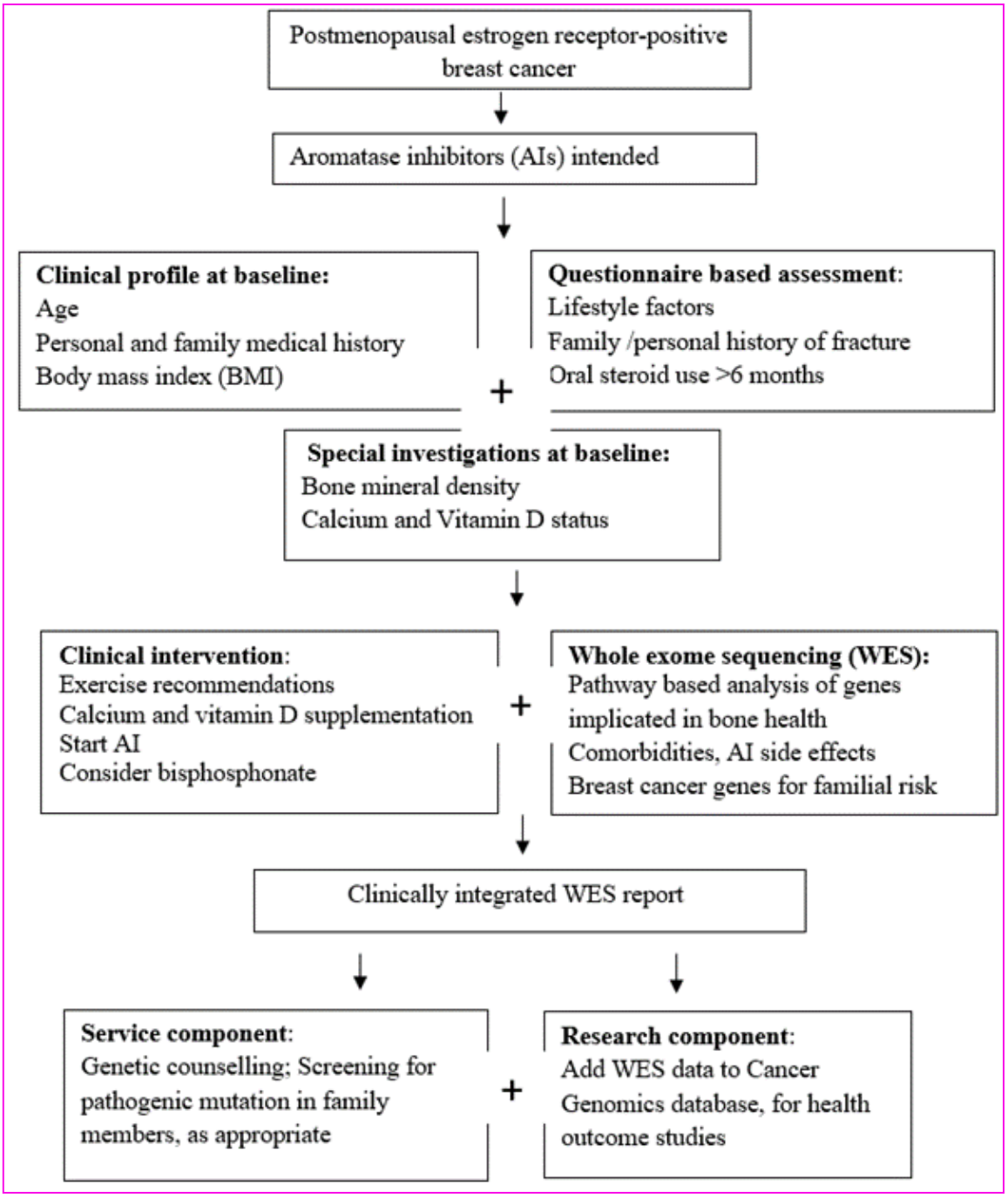


Figure 2.3: Utilisation of a pathology-supported genetic testing platform for incorporation of research-generated whole exome sequencing results into an adaptable patient report compiled from the genomics database (Source: Baatjes 2018, reproduced with permission).

2.12 Vitamin D-related genes and breast cancer

Several genes control the vitamin D pathway, therefore common genetic variants may play a role in the differences in vitamin D status among individuals and populations. From a GWAS study, determination of the influence of vitamin D pathway SNVs may add to the identification of more reliable variants of clinical significance that may serve as drug targets (Berry & Hyppönen, 2011). Lu and colleagues (2012) examined 6 common variants in 4 genes *(GC-rs4588, rs7041, rs2282679, rs1155563; NADSYN/DHCR7-rs3829251, rs1790349; CYP2R1-rs2060793)* in a population-based cohort of 3,210 mixed Chinese Han population from Beijing and Shanghai. They observed different genetic variants which might be linked to serum vitamin D levels and concluded that the *GC* and *NADSYN1/DHCR7* loci individually and collectively are associated with variation in plasma vitamin D levels in Chinese Han. This study showed that diverse genetic backgrounds might exist between these two subpopulations (Lu et al., 2012). Zhang and colleagues (2013) also investigated the genetic variants affecting vitamin D levels in a Chinese population. They selected 96 common variants in 15 genes that regulate the vitamin D metabolic pathway in 2,897 healthy Chinese subjects from Shanghai. Their findings identified genetic variation in *CYP2R1, GC,* and *DHCR7/NADSYN1* as potential contributors to the variation in serum vitamin D levels in this population. **Table 2.5** summarizes several studies conducted on vitamin D-related SNVs associated with breast cancer subtypes in different populations. The *VDR* gene implicated in bone metastasis of breast cancer (Jeon & Shin, 2018), was also strongly implicated in osteoporosis risk (**Table 2.6**). An in-depth understanding of the genetic determinants of vitamin D status may benefit targeted treatment interventions (Signorello et al., 2011). In this context, PSGT incorporating WES proved valuable to facilitate the move from single- to multigene testing (Kotze, 2016).

Table 2.5: Summary of vitamin D-related variants studied extensively in relation to breast cancer risk

Title of the study	Ethnicity	Aim	Design	Genes/SNVs studied	Reference
Genetic variation in vitamin D-related genes and risk of breast cancer among women of European and East Asian descent	Women of European and East Asian descent	Investigated associations of vitamin D-related polymorphisms (*GC* and *VDR*) and breast cancer risk and potential interactions with menopausal status and breast tumour subtypes.	Case/control 1037/1050	*VDR (rs4328262, rs11168292, rs1544410, rs7967152 and rs2239186), GC (rs7041).*	(J. Shi et al., 2016)
Association between the *BsmI* Polymorphism in the Vitamin D Receptor Gene and Breast Cancer Risk: Results from a Pakistani Case- Control Study	Pakistani women (Asian)	Investigated the possible contribution of *SNVs rs1544410 (BsmI)* and *rs2228570 (FokI)* to breast cancer risk.	Case/control 463/1012	*BsmI, FokI*	(Usman Rashid et al., 2015)
Vitamin D receptor gene polymorphisms and breast cancer risk among postmenopausal Egyptian women	Egyptian women (African)	To investigate the role of VDR polymorphisms *BsmI (rs1544410), ApaI (rs7975232), TaqI (rs731236), and FokI (rs 10735810)* in pathogenesis of breast cancer.	Case/control 130/100	*BsmI, ApaI, FokI and TaqI*	(Abd-Elkader Abd-Elsalam et al., 2015)
Single-Nucleotide Polymorphisms in Vitamin D–Related Genes May Modify Vitamin D–Breast Cancer Associations	United States or Puerto Rico residents	Examined effect modification by SNPs in vitamin D–related genes (*CYP24A1, CYP27B1, CYP2R1, GC, DHCR7/NADSYN1, RXRA, and VDR*).	Case/cohort 1524/1810	*VDR (rs4328262), CYP2R1 RXR*	(O'brien et al., 2017)
Association between vitamin D receptor gene Cdx2 polymorphism and breast cancer susceptibility- A meta-analysis of four case/control studies	Caucasian, Asian, and African	Estimation of the association between *VDR Cdx2* polymorphism and breast cancer susceptibility.	Case/control (3841/5039)	*VDR Cdx2*	(Zhou et al., 2013)
Vitamin D Receptor *Cdx-2* Polymorphism and Premenopausal Breast Cancer Risk in Southern Pakistani Patients	Pakistani Patients (Asian)	Examined the association between *VDR-Cdx2* polymorphism and breast cancer in premenopausal females.	Case/control 103/161	*VDR Cdx2.*	(un Nisa Iqbal et al., 2015)

Association between Vitamin D receptor *(Cdx2, Fok1, Bsm1, Apa1, Bgl1, Taq1, and Poly (A))* gene polymorphism and breast cancer: A systematic review and meta-analysis	Asian, Caucasian, African, American, Hispanic, Hawaiian.	Examined the relationship between VDR gene polymorphisms and breast cancer.	Case/control (26372/32883)	*VDR, Bsm1, Fok1, Apa1, Poly (A), Cdx2, Bgl1, Taq1*	(Mehir un Nisa Iqbal and Taseer Ahmed Khan, 2017)
Vitamin D-related gene polymorphisms, plasma 25-hydroxyvitamin D, and breast cancer risk	Caucasian, African and Other	Examined the associations between vitamin D-related genetic polymorphisms *(CYP27B1, VDR, CYP24A1* and *GC)* plasma 25-vitamin D, and breast cancer risk.	Case/controls 967/993	*CYP24A1 (rs6068816), rs13038432, rs3787557, rs927650), VDR (rs731236)*	(Reimers et al., 2015)
Association of low penetrance vitamin D receptor *Tru9I (rs757343)* gene polymorphism with risk of premenopausal breast cancer	Pakistani women (Asian)	Determined whether a novel polymorphism (*Tru9I*) in the low penetrance vitamin D receptor (*VDR*) gene is associated with risk of premenopausal breast cancer (BC).	Case/controls 228/503	*VDR Tru9I*	(un Nisa Iqbal et al., 2018)
Vitamin D Receptor Gene Polymorphism: Association with Susceptibility to Early-Onset Breast Cancer in Iranian, *BRCA1/2*-Mutation Carrier and non-carrier Patients	Iranian women (Asian)	Investigated *BsmI and FokI* polymorphisms in Iranian young (≤ 35 years old) breast cancer patient with known *BRCA1/2* germline mutations.	Case/controls 203/214	*BsmI, FokI*	(Shahabi et al., 2018)

Table 2.6: Vitamin D-related genes and selected variants studied extensively in relation to osteoporosis

Title of study	Ethnicity	Aim	Design	SNVs studied	Reference
Associations between VDR Gene Polymorphisms and Osteoporosis Risk and Bone Mineral Density in Postmenopausal Women	Asian and Caucasian	Determined significant association between VDR gene polymorphisms and susceptibility to osteoporosis and BMD in postmenopausal women A systematic review and Meta-Analysis	Case/control 5146/6216	*VDR ApaI, BsmI, Cdx2, FokI and TaqI*	(Zhang et al., 2018)
Association of Vitamin D Receptor Gene Variation With Osteoporosis Risk in Belarusian and Lithuanian Postmenopausal Women	Belarusian and Lithuanian women (Caucasian)	Assessed the frequency of distribution of VDR genetic variants with established effect and evaluate their haplotype association with the risk of PMO	Case/control 149/172	*rs7975232, rs1544410, rs731236,* and *rs11568820)*	(Marozik et al., 2018)
The Influence of Vitamin D Receptor Genetic Variants on Bone Mineral Density and Osteoporosis in Chinese Postmenopausal Women	Chinese Postmenopausal Women (Asian)	Evaluated the potential association between genetic variants of VDR gene and bone mineral density (BMD) and osteoporosis in Chinese postmenopausal women.	Case/control 482/488	VDR p.Gly14Ala and p.His305Gln	(He et al., 2015)
Common allelic variants of the vitamin receptor D gene *rs7975232* (ApaI) do not influence bone mineral density figures in postmenopausal osteoporotic women	Spanish postmenopausal women (Hispanic)	Investigated the relationship of commonly studied polymorphisms in the VDR gene, *rs7975232,* with the BMD figures in a cohort of Spanish postmenopausal women.	Stratified Case 274 postmenopausal osteoporotic Spanish women	*rs7975232*	(Pedrera-Canal et al., 2015)

2.13 Next generation sequencing

Sanger and colleagues developed a chain termination method for DNA sequencing (Sanger et al., 1977), which was later renamed as Sanger sequencing. Over the years, sequencing technologies have been improved with development of various NGS technology platforms **(Table 2.7),** which have transformed biological and biomedical research (Mardis, 2017). NGS capacity to sequence DNA at increasing throughput and decreasing cost (Mardis, 2017) has facilitated the use of sequencing as a clinical tool (Goodwin et al., 2016).

Table 2.7: Characteristic weaknesses and strengths of sequencing platforms relevant to this study

Platform	Read length (bp)	Weakness	Strength
Sanger sequencing			
ABI 3500/3730	Up to 1 kb	Cost and throughput	Read accuracy and length
Ion Torrent			
Proton	Up to 400 bp	Homopolymers	Speed, throughput
Oxford Nanopore			
MinION	Up to 100 kb	High error rate, run length	Read length, portability

Modified from: Next-Generation Sequencing Technologies and their Application to the Study and Control of Bacterial Infections. (Besser et al 2018).

NGS technology platforms perform sequencing of millions of small fragments of DNA in parallel, followed by bioinformatics analyses to obtain clinically useful information by mapping the individual reads to a human reference genome (Behjati & Tarpey, 2013). About three billion bases in the human genome are sequenced multiple times, which provides high depth to deliver accurate data and insight into DNA variation applicable to clinical practice. The advantages of NGS include the following abilities:

1) To capture a broader spectrum of variants than Sanger sequencing
2) To compare the results with reference genomes without bias (Dawson et al., 2013; Cristofanilli & Fortina, 2013; Behjati & Tarpey, 2013).

With these advantages, application of NGS have improved cancer genome projects around the world, providing a guide to a more precise diagnosis and classification of the disease. This led to more accurate prognosis and identification of causal mutations used as targets for effective drug treatment (Behjati & Tarpey, 2013). Individual cancer sequencing of germline DNA and tumour genetics provide the basis for personalised cancer management (Behjati & Tarpey, 2013).

The key limitation of NGS is the high cost of setting up the platform, including computer capacity, information storage and the personnel expertise needed to comprehensively analyse and interpret the data generated. Proper management and understanding of data generated using sophisticated bioinformatics tools to extract clinically important information from whole genomes in a clear and robust interface, is another challenge facing NGS (Behjati & Tarpey, 2013). Computational tools to predict the outcomes of NGS data analysis have limited capacities to do so (Li et al., 2016) because they are still evolving. NGS can be used to select gene panels or the coding regions of all ~20,000 genes in the human genome (exomes) (Behjati & Tarpey, 2013; Hodges et al., 2007).

2.13.1 Whole exome sequencing

WES is a method for identifying protein coding variants (missense, non-sense, splice site, and small deletion or insertion mutations) to produce sequence data from hundreds of millions of short DNA fragments in parallel (Singleton, 2011). WES has been used to successfully detect variants of clinical importance in patients with rare Mendelian disorders (Stray-Pedersen et al., 2017) and to study complex diseases such as breast cancer (Bai et al., 2014). The first step in WES is the creation of a DNA library that consists of enriched targeted DNA regions of interest. In the context of exome sequencing, this target selection is performed with amplification-based products, manufactured to yield a DNA sample that consists of the protein-coding and regulatory regions of the genome. **Figures 2.4** shows an overview of the steps involved in WES applied to detect variants of clinical significance for patient management.

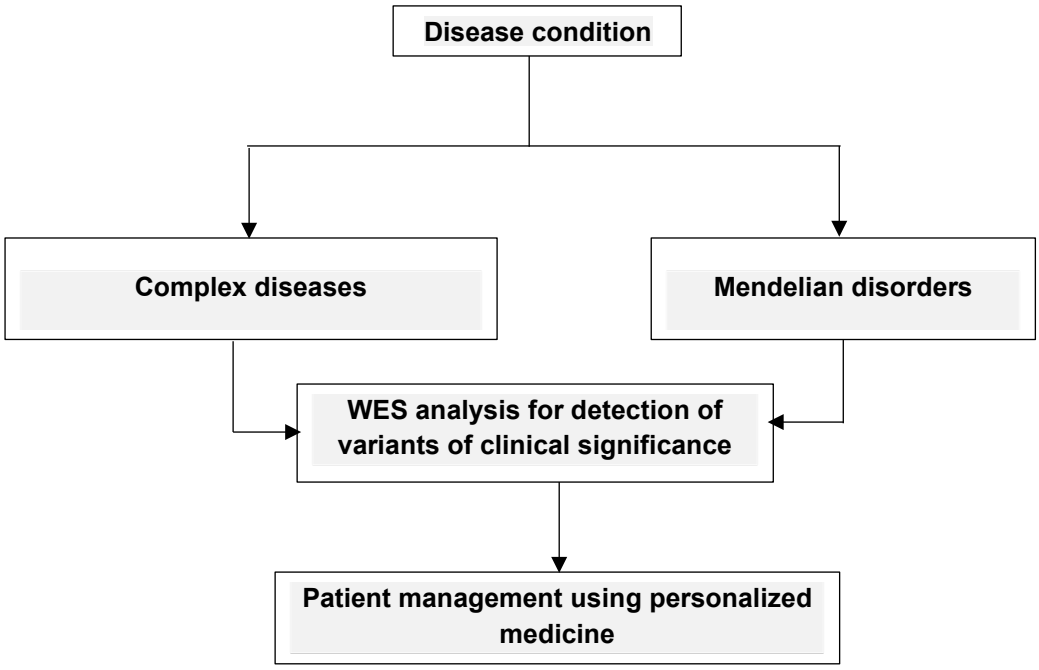


Figure 2.4: Application of WES to detect mutations of clinical significance in both Mendelian and heterogeneous diseases for improved clinical management.

The Illumina instruments are the single most commonly used of the NGS platforms and hence may lead to bias in identifying variants of clinical importance (Chakravorty & Hegde, 2017). Comparative studies of NGS platforms revealed that data generated with the Ion Torrent platform has a higher raw error rate of about 1.8% when compared with Illumina data with an error rate of about 0.4%, on the assumption of sufficient coverage (Quail et al., 2012). However, the representation and ability to call SNVs is closely matched between these technologies. The Ion Torrent platform illustrated in **Figure 2.5** applied in the present study has previously been shown to produce excellent result (Ashktorab et al., 2016).

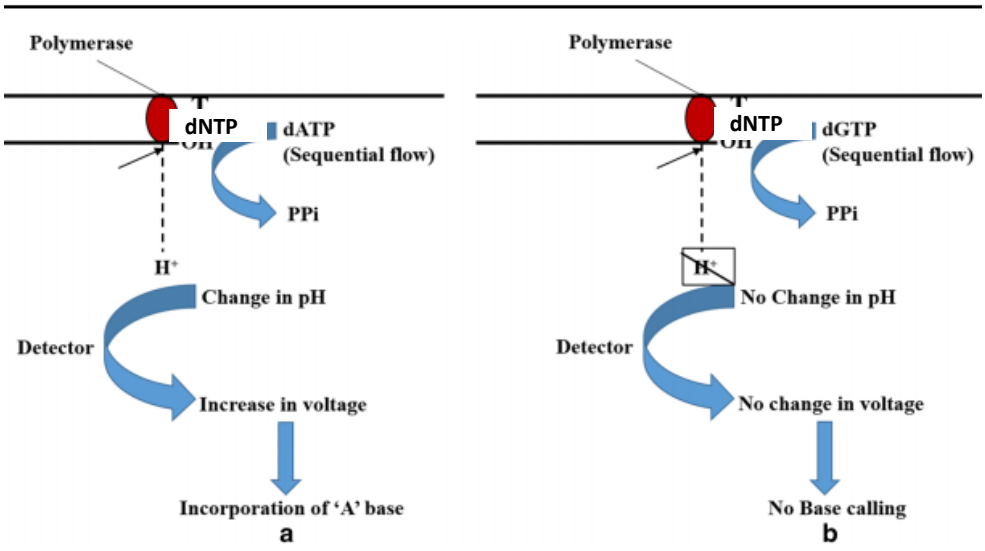


Figure 2.5: Diagrammatic representation of pH change involved in Sequencing by detection of hydrogen ions a) during complementary nucleotide incorporation and b) when a nucleotide is not incorporated. Adapted from: High Throughput Sequencing: An Overview of Sequencing Chemistry (Ambardar et al., 2016).

More recently Oxford nanopore Technologies (ONT) introduced the MinION nanopore sequencer (van Dijk et al., 2018). The MinION is a small device weighing about 87 g which can be connected to a computer through a standard USB 3.0 port (Patel et al., 2018). The development of this third generation sequencing platform paved the way for a real-time sequencing process, with generation of long reads when compared with the NGS platforms (van Dijk et al., 2018). The mechanism of nanopore sequencing is based on an electrically resistant synthetic membrane and a voltage which is applied across the membrane. Library preparation of DNA molecules are based on standard protocols initially attaching a leader adaptor and motor protein to one strand of DNA. During sequencing, the motor protein opens dsDNA, thereby allowing for passage of a base of single strand through the pore at a time (Leggett & Clark, 2017). The presence of the DNA molecule in the pore generates a deflection in the current, which is directly related to the base present in the pore at that point in time.

Improved base-calling and refining of the variant curation software used with the MinION device placed nanopore technology at the forefront of cost-effective clinical sequencing on the African continent.

2.14 Variant classification

A classification system of five categories shown in **Table 2.8** was proposed for interpretation of gene variants identified by sequencing technologies, with a goal to add more evidence in patient reports, as appropriate for individual circumstances. These include results obtained with segregation studies in families, tumour pathology, and functional *in vitro* assays with continued improvement as research data accumulate (Plon et al., 2008).

Table 2.8: Proposed Classification System for Sequence Variants Identified by Genetic testing (Adapted from: Plon et al., 2008)

Class	Description	Probability of being pathogenic
5	Definitely pathogenic	>99%
4	Likely pathogenic	>90%
3	Uncertain clinical significance	5-90%
2	Likely benign	<5%
1	Benign	<1%

Tumour histopathology is of particular importance in breast cancer patients with rare germline mutations in the *CDH1* gene. Pathogenic mutations in this gene are associated with ILC, but not invasive carcinoma of no special type (ICNST). However, tumour histopathology is currently not taken into account in variant classification calculators (Nykamp et al., 2017; Richards et al., 2015) due to increased complexity during the variant interpretation process. This represents a knowledge gap addressed by PSGT in this study using WES to explore the role of vitamin D deficiency and *VDR* variants in postmenopausal hormone receptor-positive breast cancer patients screened for pathogenic mutations in the *BRCA1/2, CDH1* and other high- to moderate-penetrance breast cancer susceptibility genes. Pathogenic germline *CDH1* mutations predominate in families with gastric cancer (Corso et al., 2016; Sporle et al., 2018) and therefore family history is also a very important consideration for classification of variants detected in this gene. In the past a family history of cancer was mostly used to select patients

for *BRCA* mutation detection and family screening, while it now carries a new meaning in making sense of VUS frequently uncovered using advanced NGS technologies. The same applies to environmental triggers of genetic risk factors such as unhealthy lifestyle habits or medication associated with breast cancer co-morbidities and drug side effects or failure.

TNBC is associated with an increased risk for inheritance of pathogenic germline mutations in the *BRCA1* gene, while *BRCA2* mutation-positive breast cancer patients may share similar pathologic characteristics with non-carriers (Chen et al., 2018). *BRCA1* expression is critical for mediating the beneficial biological impact of vitamin D on breast cancer growth (Pickholtz et al., 2014). Vitamin D-based therapies target dormant cells that are often resistant to conventional chemotherapy regimens and may benefit patients who express some level of *BRCA1*. Thus, the genetic background of individual patients relating to high-penetrance genes such as *BRCA1/2* and *CDH1* may need to be considered to achieve effective vitamin D-based therapies or prevention. This is of particular relevance in the South African population with an increased risk for familial breast cancer due to a founder effect. This phenomenon relates to the unique genetic structure of certain South African populations which enabled the development of cost-effective genetic tests (Kotze, 2016). A small number of highly penetrant founder mutations in the *BRCA1 and BRCA2* tumour suppressor genes shown to cause early-onset familial breast cancer in a large proportion of South African patients studied at Tygerberg Hospital (Schoeman et al., 2013) has recently been incorporated into a cost-saving point of care (POC) *BRCA* assay. The previously-described Hybeacon technology (Blackman et al., 2015) was applied that allows genetic testing within 1 hour (Kotze et al., in press). By taking advantage of both the founder- and pleiotropic effects associated with *BRCA1/2* mutations (Smith et al., 2013), rapid POC DNA testing can now be performed as part of the PSGT algorithm previously described to facilitate differential diagnosis of inherited and lifestyle-related breast cancer (van der Merwe et al., 2017). When used as a first-line test in patients referred for *BRCA1/2* mutation testing implemented during a genetic counselling session, POC testing may provide a definitive diagnostic result or indicate the need for extended testing

using WES as illustrated in **Figure 2.6**. This PSGT approach currently focuses on differential diagnosis and clinical outcome of hereditary conditions related to five metabolic pathways: (i) lipid and lipoprotein metabolism; (ii) DNA methylation and mismatch repair; (iii) haemostasis and inherited thrombophilia; (iv) haem synthesis and iron homeostasis; and (v) drug metabolism (M. Kotze, 2016). Addition of the vitamin D pathway is yet to be included based on the outcome of the present study.

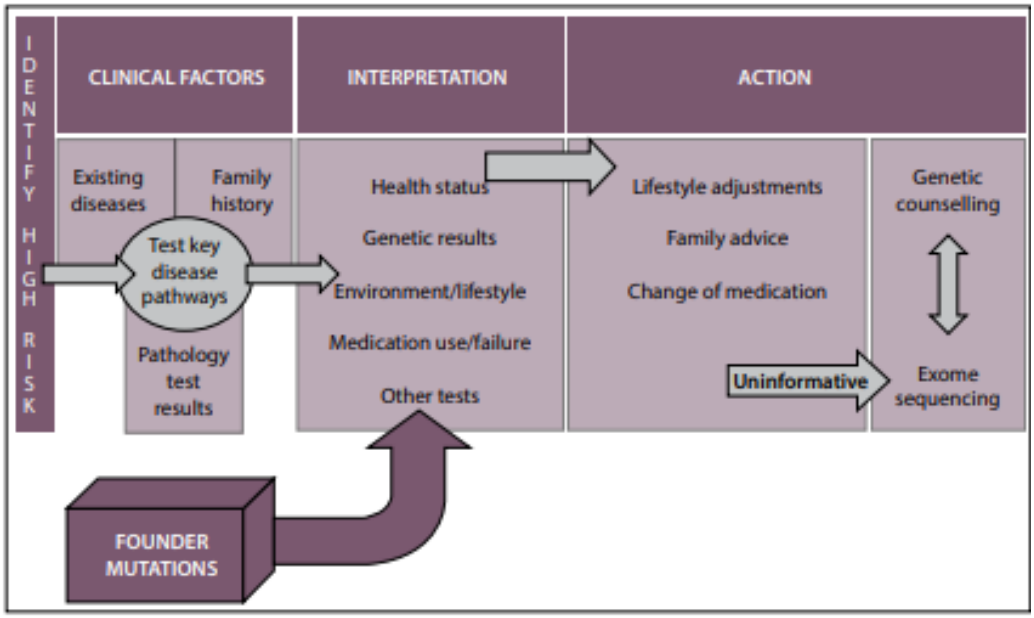


Figure 2.6: Pathology-supported genetic testing framework incorporating founder mutations and pleiotropic effects in the pre-screen algorithm for WES. (Source: Kotze 2016) with permission.

CHAPTER 3

SUBJECTS AND METHODS

3.1 Ethics approval

Ethics approval for this study was obtained from the Human Research Ethics Committee of Stellenbosch University with reference number S17/10/219. Written informed consent was obtained from all study participants prior to sample collection and electronic data capturing. This study was performed in accordance to the principles of the Declaration of Helsinki regarding human experimentation.

3.2 Study population

Only women with histologically confirmed postmenopausal hormone receptor-positive breast cancer were included in this study using the clinical criteria and methodology previously described by Baatjes et al. (2019). Ethnic distribution reflected the catchment area of the Tygerberg Academic Hospital Breast Clinic, in the Western Cape Province of South Africa. All patients were newly diagnosed with stage 0-III hormone-receptor-positive breast carcinoma at the National Health Laboratory Service (NHLS).

Inclusion criteria: Postmenopausal women aged 50-80 years with amenorrhoea for more than one year and/or a follicle-stimulating hormone (FSH) level of >40 mIU/mL; newly diagnosed with breast carcinoma and informed consent provided.

Exclusion criteria: Women with breast cancer; receiving tamoxifen therapy for breast carcinoma (currently or previously); undergoing palliative aromatase inhibitor treatment; patients known to suffer from a disease or taking drugs which adversely influence skeletal health; use of bone-protective medication (e.g. bisphosphonates) or informed consent for laboratory and genetic testing not provided.

A total of 116 participants comprising multi-ethnic groups of predominantly mixed ancestry from 126 included in the database qualified for this study based on the selection criteria shown in **Figure 3.1.**

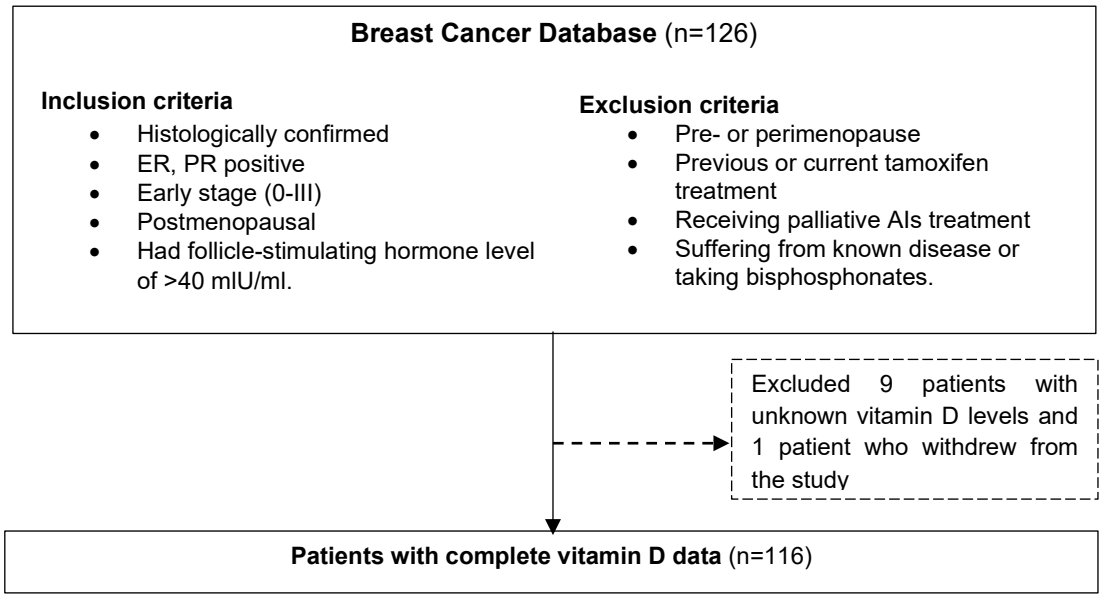


Figure 3.1: Schematic representation of the criteria used to select postmenopausal breast cancer patients from the research database.

3.3 Database

Information of breast cancer patients with previously-determined vitamin D levels was extracted from the REDCap research database developed under ethics approved project reference number S13/05/103. DNA samples linked to this database were stored in the Pathology Research Facility of Stellenbosch University that provides the infrastructure for return of research results under reference number N09/08/224. Interpretation of genetic findings in this study was performed in conjunction with a questionnaire-based health assessment and co-morbidities identified at baseline as described by Baatjes et al. (2019), using the PSGT strategy (Kotze et al. 2015). This approach involves the establishment of a genomics database at the interface between the laboratory and clinical practice with a goal to facilitating the move from single- to multi-gene testing and WES (Kotze, 2016).

3.4 Determination of vitamin D levels

The retrieved vitamin D levels of all the patients in this study were determined quantitatively at PathCare laboratory using a chemiluminescent immunoassay on the Beckman Coulter® instrument (95350 Lakeview Parkway S Drive Indianapolis, IN 46268 United States). This method is a two-step competitive binding immunoenzymatic assay. In the first step, the sample reacts with paramagnetic particles that are coated with sheep monoclonal anti-vitamin D antibodies and an agent that releases vitamin D binding protein (DBP) and the vitamin D in the sample binds to it. Alkaline phosphatase conjugate is then added to compete with binding to the immobilized monoclonal anti-vitamin D. The materials bound to the solid-phase are then held in a magnetic field while unbound materials are washed away. Chemiluminescent substrate Lumi-Phos* 530 is added and light generated by the reaction. The vitamin D concentration in the substrate was measured with a luminometer. Standard reference values were used to categorise patients with vitamin deficiency (<20 ng/mL), insufficiency (20 − 29 ng/mL) and sufficiency (>30 ng/mL) (Holick, 2009) as previously described in our multi-ethnic study population (Baatjes et al., 2019). The coefficient of variance at various clinically significant cut-off points for vitamin D levels was less than 8%. Seasons in which blood samples were collected from the patients, bone mineral density, tumour size, tumour type, tumour grade, ER, PR and HER2 were also extracted from the database for analysis.

3.5 DNA Extraction and quality control

DNA was extracted from whole blood using the QIAGEN QIAamp® DNA Blood Mini Kit (Hilden, Germany). The NanoDrop One Spectrophotometer (Thermo Electron Scientific Instruments LLC, Madison, WI, USA) was used to measure the DNA quantity. All genomic DNA samples were diluted to a final concentration of 10 ng/µL using nuclease-free water.

The quality of the DNA used for WES was evaluated on the NanoDrop™ ND-1000 spectrophotometer (ThermoFisher Scientific), using a low Tris-EDTA buffer. The double

stranded DNA (dsDNA) fraction in each sample was determined on the Qubit 4.0 fluorometer (ThermoFisher Scientific) using the Qubit™ dsDNA HS assay kit. Genomic quality scores (GQS) were determined on the LabChip® GXII Touch (PerkinElmer, Waltham, MA, USA), using the DNA Extended Range Chip and genomic DNA (gDNA) Reagent Kit (PerkinElmer). The GQS is a value between 0 and 5, with zero indicating the lowest quality DNA and five high-quality, intact gDNA.

3.6 Whole exome sequencing

WES was performed using the Ion Torrent apparatus at the Central Analytical Facility of Stellenbosch University. DNA samples of patients selected for WES were screened for genetic variation in the genes listed in **Table 3.1**, including *BRCA1/*2 routinely screened for pathogenic mutations at Tygerberg Academic Hospital (Schoeman et al., 2013; Lopes et al., 2012; Lee et al., 2016; Cabanillas et al., 2017). The WES results of seven samples were compared with variants detected in the *CDH1* gene using nanopore long-range sequencing.

Table 3.1: Genes screened for clinically relevant variants in the study population using whole exome sequencing

High penetrance	**Moderate /variable penetrance**				**Low penetrance**
BRCA1 NM_007294.3	*ATM* NM_000051.3	*BARD1* NM_000465.3	*MSH6* NM_000179.2	*RAD51B* NM_002877.5	*GC* NM_001204307
BRCA2 NM_000059.3.1	*BRIP1* NM_032043.2	*EPCAM* NM_002354.2	*NBN* NM_001024688.2	*RAD51C* NM_002876.3	*VDR* NM_001017536
CDH1 NM_001317184.1	*CHEK2* NM_007194.3	*FANCC* NM_000136.2	*NF1* NM_001042492.2	*RAD51D* NM_002878.3	*DHCR7* NM_001360
PTEN NM_000314.6	*PALB2* NM_024675.3	*MLH1* NM_000249.3	*PMS1* NM_000534.4	*RECQL* NM_002907.3	*CYP2R1* NM_024514
STK11 NM_000455.4		*MRE11A* NM_005590.3	*PMS2* NM_000535.6	*MUTYH* NM_001048171.1	*CYP27B1* NM_000785.4
TP53 NM_001126114.2		*MSH2* NM_000251.2	*RAD50* NM_005732.3	*XRCC2* NM_005431.1	*CYP24A1* NM_001128915

3.6.1 Ion Torrent platform

The Ion AmpliSeq™ Exome RDY Kit was used to prepare whole human exome libraries from 100 ng input gDNA. Targets were amplified on the SimplyAmp Thermal Cycler using the Ion AmpliSeq™ Exome RDY Panel and the Ion AmpliSeq™ Library Kit Plus. Following exome target amplification across primer pairs, the products were combined, and primer sequences partially digested in preparation for adapter ligation. IonCode™ Barcode Adapters were used to generate barcoded libraries and adapter ligated exome libraries were purified with Agencourt™ AMPure™ XP reagent and eluted in 50 µl low TE buffer. Figure 3.2 is the schematic representation of the process. Exome library quantification was performed using the Ion Library TaqMan Quantitation Kit (MAN0015802 Rev B.0). Briefly, qPCR reaction volumes were scaled to a final volume of 10 µl and amplification was performed using the StepOnePlus™ Real-time PCR system (ThermoFisher Scientific). Libraries were diluted to a target concentration of 20 pM. The diluted, barcoded exome libraries were combined in equimolar amounts for sequencing template preparation using the Ion 550™ Chef Kit. Twenty-five microliters of diluted, pooled library was loaded onto the Ion Chef liquid handler for template preparation and enrichment using Ion 550™ Chef Reagents, Solutions and Supplies. Enriched ion sphere particles were loaded onto an Ion 550™ Chip. Sequencing was performed on the Ion S5™ platform.

3.6.2 WES bioinformatics analysis

Flow space calibration, base caller analysis, reference genome alignment to a synthetic Caucasian major allele reference sequence (CEU-MARS), coverage analysis and variant calling were performed **(Figure 3.2)** as previously described by van der Merwe et al., 2017. These procedures involve alignment of the generated sequence reads to a Major Allele Reference Sequence (MARS) (Dewey et al., 2011). The torrent alignment program (TMap)

was used for the alignment of short and long nucleotide sequences produced by NGS (https://github.com/iontorrent/TS/tree/master/ Analysis/TMAP).

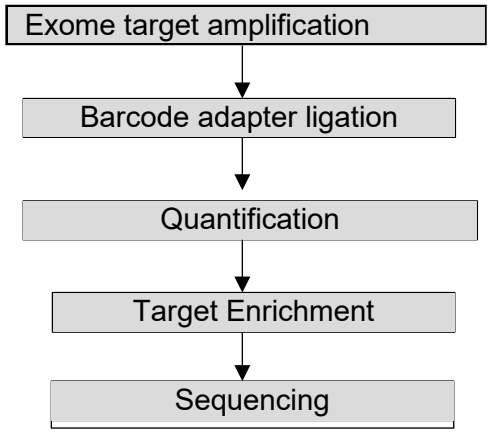


Figure 3.2: Summary of Ion AmpliSeq Exome RDY workflow

The second step involved read analysis and processing was performed using the preinstalled Torrent Suite™ Software and variant calling with the Torrent Variant Caller (TVC) plug-in. Mapped reads were processed in BAM file format as shown in **Figure 3.3**.

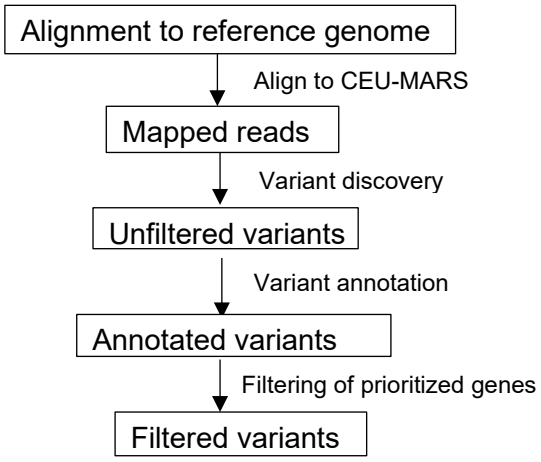


Figure 3.3 Schematic representation of sequencing read processing pipeline

A minimum Phred quality score (Q score) of 50 was used for base calling, allowing interpretation with about 99% accurate base call. Higher Q scores correspond to higher

quality. A minimum depth of coverage of 100 was used for confident variant calls. Coverage provides counts of read depth (DP) at two different levels: at 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. Resulting variants were assessed in relation to publicly available mutation and variant databases. These include Clinvar (http://www.ncbi.nlm.nih.gov/clinvar/), 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) which obtains information from PubMed that is regularly updated to support genome annotation and interpretation prior to analytical validation of potentially deleterious gene variants in the laboratory. The Varsome calculator (www.varsome.com), incorporating functionality scores from bioinformatics tools such as Polymorphism Phenotyping v2 (PolyPhen-2), the Sorting Intolerant from Tolerant (SIFT) algorithm and the Protein Variation Effect Analyzer (PROVEAN), was used to determine the effect of variants on the biological function of resulting proteins. Finally, all potentially significant variants detected were verified with the Integrative Genomics Viewer (IGV) - a high-performance visualization tool used for interactive exploration of large, integrated data sets (www.broadinstitute.org/igv/). Allele frequencies and genotype distributions of the SNVs detected were compared with minor allele frequencies (MAF) reported for different populations in the ExAC Database (exac.broadinstitute.org).

3.7 Real-time polymerase chain reaction

The LightCycler LC480-II instrument (Roche Applied Science, Penzberg, Germany) was used for real-time PCR. The ABI™ TaqMan® SNP Genotyping assay ID C_8278879_10 was used according to the standard TaqMan® SNP genotyping protocol. Pre-designed 40× assay mix consisted of unlabelled primers and TaqMan® Minor Groove Binder (MGB) probes (VIC® and FAM™ dye-labelled) used to differentiate between the three genotypes of *GC rs4588*. Two

microliters of template DNA (10 ng/µL) was used in a total reaction volume of 10 µL with RealQ Plus Master Mix for Probe (Ampliqon, Odense, Denmark).

3.8 Sanger sequencing

PCR-based Sanger sequencing was used as the gold standard for confirmation of potential clinically relevant gene variants and to confirm correct allocation of SNV genotypes (homozygosity and heterozygosity). Oligonucleotide primers were designed with PrimerBLAST to sequence the DNA regions of interest by Sanger sequencing of the PCR products. Oligonucleotide sequences were verified on reference sequences obtained from Ensembl (ensembl.org). **Table 3.2** shows the oligonucleotide primers used for Sanger sequencing.

Table 3.2: Oligonucleotide primers used for conventional PCR and Sanger sequencing

Gene	dbSNP ID	Primer	Oligonucleotide Primers	Size (bp)
GC	*rs4588*	Forward	TCTCGAAGAGGCATGTTTCAC	321
		Reverse	TCACAGTAAAGAGGAGGTGAGT	
VDR	*rs7975232 (ApaI)*	Forward	ACGTGGTCTGGGCTACAG	655
		Reverse	CACTCAGGCTGGAAGGAGAG	
VDR	*rs1544410 (BsmI)*	Forward	AAGCTGAACTTGCATGAG	532
		Reverse	CTTCTCACCTCTAACCAG	
VDR	*rs2228570 (FokI)*	Forward	CAGCTATGTAGGGCGAATC	339
		Reverse	CAGCCTTCACAGGTCATAGC	
VDR	*rs731236 (TaqI)*	Forward	TCACCGGTCAGCAGTCATAG	388
		Reverse	CATTGCCAAACACTTCGAG	
CDH1	*rs201511530*	Forward	TGGGATCCTTCTTTACTAATTCTTTTTCTTTCA	232
		Reverse	TCTCTTAGAAGCTTGTTGACACCG	

Confirmation of successful PCR amplification was performed by adding 1 µL of 500x SYBR Green fluorescent dye to the PCR tube after thermal cycling and checking for strong green fluorescence under UV light. PCR products were submitted to the DNA sequencing Central

Analytical Facility (CAF) of Stellenbosch University for post-PCR clean-up and automated sequencing. Electropherograms were analyzed using FinchTV Version 1.4.0 software (Geospiza Research Team). With this application, the nucleotide sequences of each gene were compared directly to the Ensembl human reference sequence.

3.9 Long-range nanopore sequencing (MinION platform)

Long-range PCR (LR-PCR) and nanopore sequencing of overlapping amplicons covering the entire *CDH1* gene was performed using the pocket-sized MinION device (Oxford Nanopore Technologies). After base calling with Guppy and mapping with minimap2, variants were called with Varscan2. The entire *CDH1* gene was covered in 6 amplicons **(Table 3.3)** per sample by LR-PCR performed with PrimeStar GXL polymerase (Takara Bio Inc.), according to the manufacturers' protocol.

Table 3.3: Oligonucleotide primers used for long range MinION sequencing of CDH1 gene

Gene	dbSNP ID	Primers	Oligonucleotide Primers for MinION sequencing	Size (bp)
CDH1	*rs376097289*	Forward Reverse	ACTCCAGGCTAGAGGGTCAC TTTCCAACCCCTCCCTACTC	1350
CDH1	*rs376097289*	Forward Reverse	AGCGTTCAATTCCCCTGCT CCCGAACCCTTTCTCACACA	8036
CDH1	*rs376097289*	Forward Reverse	TCTCTGTGATTTCTGCCCTGC TAACAACCCCAAACTGTCCCA	7380
CDH1	*rs376097289*	Forward Reverse	GGAAAAGACCCAGTGTTGGGAT GCAAGTCAGTTGAAAAATCCTCAC	7158
CDH1	*rs376097289*	Forward Reverse	CCAGAGCTTGTCCCCGTTCAG AGTCCCCATCCATGCAAAGC	7525
CDH1	*rs376097289*	Forward Reverse	TGGCTCTCAACACTTGCTCT CTCTTGACTGGAAATGTGGTGG	1350

Eppendorf DNA Lo-Bind tubes were used throughout the entire nanopore library preparation and reaction products were purified with AMPure XP beads (Beckman Coulter Inc.) after every step in the protocol. For AMPure XP bead purification, samples were mixed with an appropriate volume of AMPure beads and incubated at room temperature on a Hula mixer for 5 minutes with vortexing. Beads were pelleted on a magnet stand and washed twice with freshly prepared 80% ethanol. After air-drying for 2 minutes, the bead pellets were resuspended in elution buffer and incubated at room temperature for 5 minutes. The tubes were then placed on the magnet stand and the beads allowed to pellet for 2 minutes, after which the eluate containing the DNA was transferred to a new Eppendorf DNA Lo-Bind tube. Each PCR product was quantified with Qubit 1.0 and amplicons were pooled per sample in equimolar amounts to a total of 1.5 µg. The *CDH1* amplicon pools were purified with AMPure XP beads and eluted in 50 µL nuclease-free water. This was followed by library preparation using the nanopore 1D native barcoding protocol (SQK-LSK109) from Oxford Nanopore Technologies. PCR products (~1.5 µg DNA) were end-repaired prior to adapter ligation in a total reaction of 60 µL, comprising 50 µL of amplicons, 7 µL of Ultra II End-Prep reaction buffer and 3 µL Ultra II End-Prep enzyme mix (New England Biolabs), and incubated in a thermocycler at the appropriate temperatures (20°, 65°) for 5 minutes each.

Aiming for DNA template >700 ng, 1 µL of the end-prepped DNA samples was quantified with a Qubit fluorometer. Next, the native barcode ligation step was carried out on the repaired and end-prepped DNA samples using the nanopore barcoding kit EXP-NBD103. The library was loaded into the SpotON sample port of a MinION flow cell R 9.4.1 according to the manufacturer's instructions.

3.9.1 MinION bioinformatics analysis

To translate the electric signals from the flowcells into raw DNA sequence reads (i.e. A, C, G and T) the raw data (fast5 file format) are first basecalled with the Guppy basecaller (Oxford Nanopore Technologies, v2.2.3). For improved speed, Guppy GPU basecalling was done on

an NVIDIA GeForce GTX1050 Ti graphics card which generated fastq files. After demultiplexing the reads according to their barcodes (also with Guppy), the fastq files were mapped to the CEU reference genome with Minimap2. Samtools was used to convert .sam files to .bam files and to sort and index the .bam files. Variants were called using VarScan software (v2.2.3). **Figure 3.3** shows a summary of the pipeline used for analysis of the data generated on the MinION device. The results of *CDH1* variants detected in 7 DNA samples were compared with WES reads analysed with the Ion Torrent Suite.

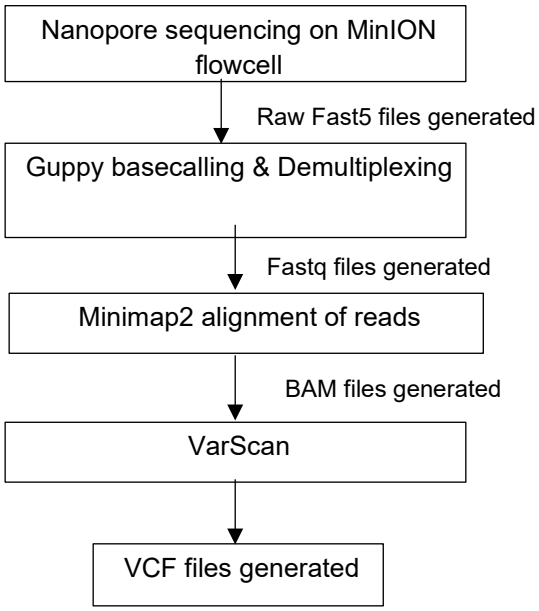


Figure 3.3: Schematic representation of sequencing read processing pipeline for MinION data.

3.10 Statistical Analysis

The statistical tools used were STATA, STATISTICA and R. For descriptive purposes, crosstabulation with the Fisher exact test was performed to investigate the relationship between categorical variables. Analysis of variance (ANOVA) was used to compare mean vitamin D status between the four seasons. Post hoc testing was performed using Fisher Least Significant Difference (LSD) testing. Normality was checked by inspecting normal probability plots and were found to be acceptable. Levene's test was used to assess homogeneity of variance assumptions. Pearson correlation was applied to test the relationship between BMI and vitamin D levels. Receiver Operating Curve (ROC) analysis was used to determine the vitamin D cut-off point for best discriminating between genotype groups. Genotype distribution and allele frequencies were determined by allele counting. Hardy-Weinberg equilibrium (HWE) was determined using an online calculator (http://www.oege.org/software/hwe-mr-calc.shtml). Clinical characteristics were compared between study subgroups of breast cancer patients to determine possible allelic effects of selected variants. Due to the exploratory nature of this study, multivariate analyses were not performed in order to avoid misinterpretation of the results due to small sample size.

CHAPTER 4

RESULTS

This study included a clinical assessment of patient information extracted from a central research database using vitamin D levels measured in hormone receptor-positive postmenopausal breast cancer patients as the primary selection criteria, followed by genetic studies including the *CDH1*, encoding E cadherin activated by the vitamin D receptor (VDR).

4.1 Clinical assessment of risk factors for vitamin D deficiency

Vitamin D levels of 116 postmenopausal women categorised as deficient (53%), insufficient (38%) or sufficient (9%) **(Figure 4.1)** were extracted from the research database for correlation studies performed in relation to the clinical characteristics of the study population summarised in **Table 4.1.**

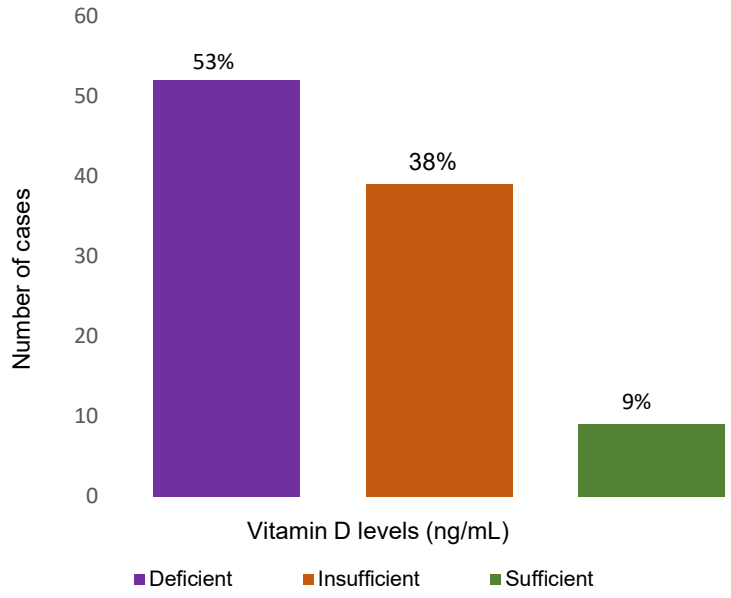


Figure 4.1: Study population of 116 postmenopausal breast cancer patients grouped according to vitamin D deficiency, insufficiency and sufficiency.

Table 4.1: Clinical characteristics of the study population

Variable	Study Group	Total (n=116)
*Body mass index (kg/m^2)	Mean ± SD	31.66 ± 7.8
	<26	30 (27%)
	26-30	22 (20%)
	>30	59 (53%)
*Physical activity	Occasionally	91 (82%)
	Regularly	20 (18%)
*Alcohol intake	Abstain	91 (82%)
	1-7 units per week	17 (15%)
	>7 units per week	3 (3%)
*Smoking	Never	73 (66%)
	Current	30 (27%)
	Previous	8 (7%)
Vitamin D (ng/mL)	Mean ±SD	20.26 ± 6.95
Tumour grade (n=90)	I	32 (31%)
	II	49 (47%)
	III	23 (22%)
Tumour type (n=111)	ICNST	83 (74%)
	ILC	14 (13%)
	Other	14 (13%)
Tumour size (mm) (n=113)	≤ 20	48 (42%)
	≥21	65 (58%)
ER status (n=116)	Positive	113 (97%)
PR status (n=116)	Positive	92 (79%)
HER2 status (n=109)	Positive (3+)	14 (13%)
*Bone mineral density (n=100)	Normal	58 (58%)
	Osteopenia and Osteoporosis	42 (42%)

*Data overlap with data previously described by Baatjes et al. 2019; SD, standard deviation; ICNST: invasive carcinoma of no special type; ILC: invasive lobular carcinoma; ER: oestrogen receptor; PR: progesterone receptor; HER2: human epidermal growth factor receptor-2; BMI: body mass index (<26 kg/m2: normal, 26-30 kg/m2: overweight, >30 kg/m2: obese)

As indicated in Table 4.1, the effect of age, physical activity, alcohol consumption, smoking and BMI were considered as contributing factors to vitamin D levels, as well as season of blood collection (**Figure 4.2**). A statistically significant effect of seasonal variation on mean

vitamin D levels was detected, with the lowest levels measured during winter, independent of BMI and ethnicity ($p < 0.01$). The negative association between BMI and vitamin D shown in **Figure 4.3** was independent of age, ethnicity and seasonal changes (r = -0.26, p<0.01).

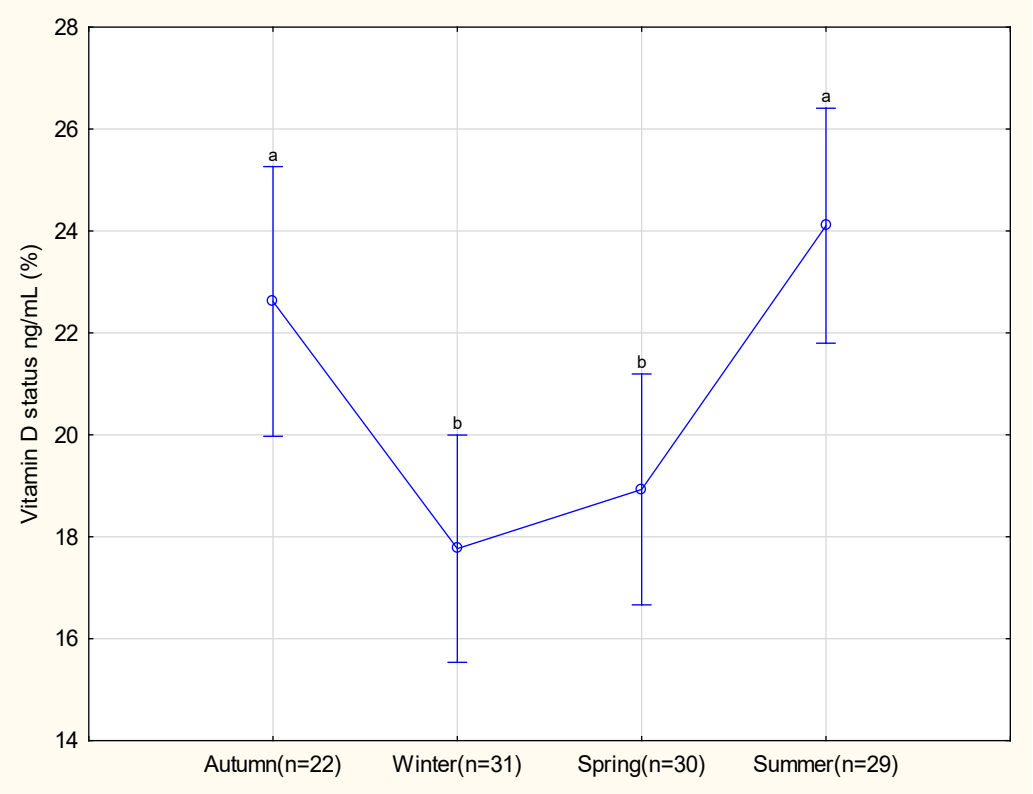


Figure 4.2: Effect of seasonal variation on vitamin D levels in the study population.

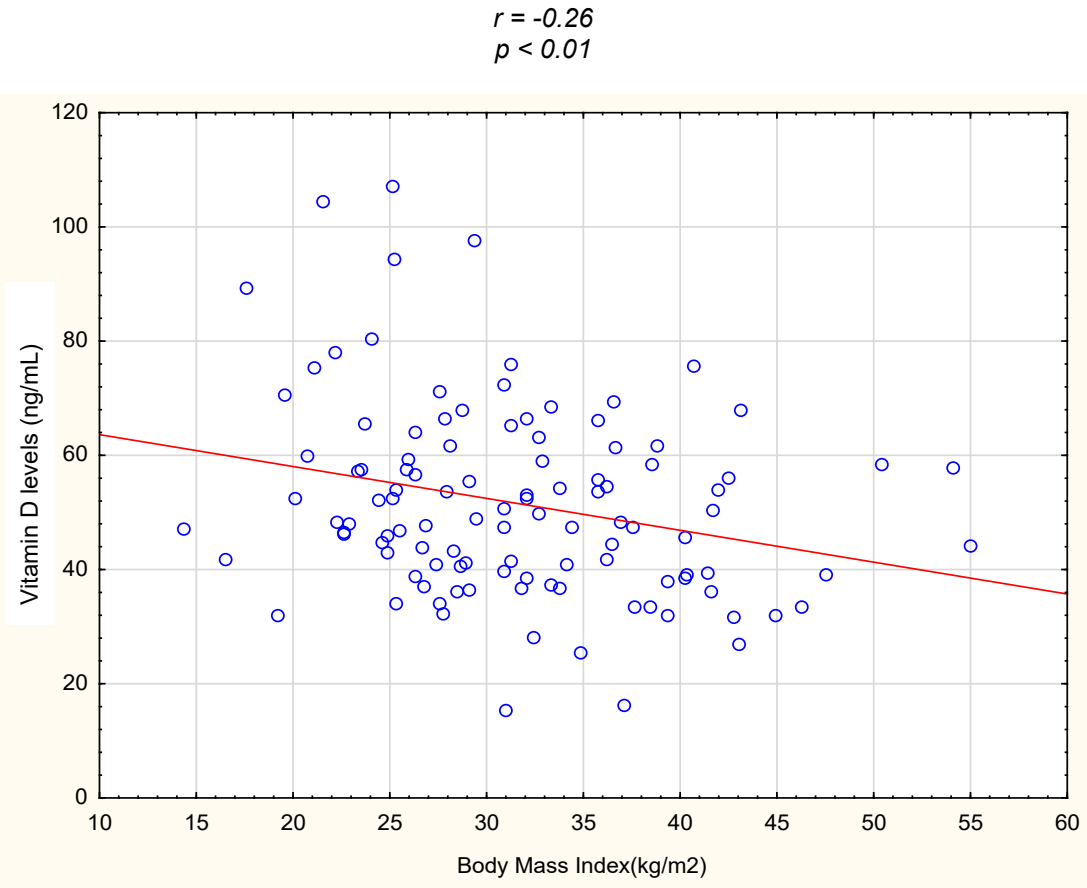


Figure 4.3: Correlation between body mass index (BMI) and vitamin D levels in the study population.

The above-mentioned findings provided the clinical context for investigation of the potential effect that vitamin D levels may, in turn, have on tumour pathology. No statistically significant association was detected between vitamin D levels and tumour type, size, grade, ER, PR and HER2 status. **Figure 4.4** shows the relationship between vitamin D levels and tumour type, indicating that none of the patients with invasive lobular carcinoma had sufficient vitamin D levels.

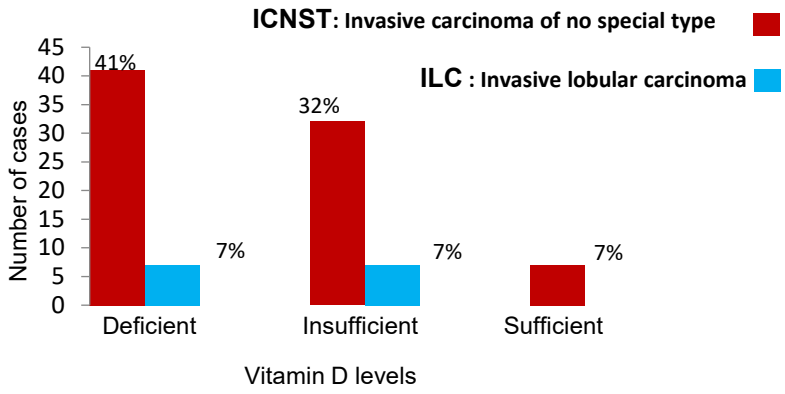


Figure 4.4: Relationship between vitamin D levels and tumour pathology based on invasive carcinoma of no special type and invasive lobular carcinoma classification.

4.2 Whole exome sequencing for identification of vitamin D-lowering genetic variants

The initial genetic screen using WES included 10 patients at the extremes of vitamin D levels (≤12 ng/mL vs >30 ng/mL). **Table 4.2** shows two SNVs detected in exon 11 (*rs4588)* and exon 12 (*rs7041)* of the *GC* gene, together with BMI and the season of blood collection for measurement of vitamin D levels. Potential clinically relevant SNVs were not detected in the other vitamin D pathway genes analysed, including *DHCR7, CYP2R1, CYP27B1* and *CYP24A1*.

Table 4.2: Missense variants detected by WES in 10 breast cancer patients with the lowest and highest extremes of vitamin D levels found to be influenced by seasonal changes and BMI in the study cohort.

SAMPLE	VITAMIN D (ng/mL)	SEASON	BMI (kg/m^2)	GENE	BASE CHANGE	SNV (Genotype)	AMINO ACID CHANGE	PATHOGENICITY SCORE (DANN Score)
004	4.82	Winter	Obese	*GC*	c.1364C>A c.1353T>G	*rs4588 (GT)* *rs7041 (GT)*	T455K D451E	0.187 0.4603
007	10.3	Spring	Obese	*GC*	c.1364C>A	*rs4588 (GT)*	T455K	0.187
051	12.8	Winter	Normal	*GC*	c.1364C>A	*rs4588 (GT)*	T455K	0.187
093	11.6	Autumn	Obese	*GC*	c.1364C>A	*rs4588 (TT)*	T455K	0.187
124	12.4	Summer	Obese	*GC*	c.1364C>A c.1353T>G	*rs4588 (GT)* *rs7041 (GT)*	T455K D451E	0.187 0.4603
073	39.2	Summer	Overweight	*GC*	c.1364C>A	*rs4588 (GT)*	T455K	0.187
077	31.3	Summer	Normal	*GC*	c.1364C>A	*rs4588 (GT)*	T455K	0.187
089	41.9	Autumn	Normal	*GC*	c.1353T>G	*rs7041(GT)*	D451E	0.4603
092	37.9	Autumn	Normal	*GC*	c.1353T>G	*rs7041(GT)*	D451E	0.4603
122	43	Spring	Normal	*GC*	c.1353T>G	*rs7041(GT)*	D451E	0.4603

BMI, body mass index; GC *rs4588,* T=0.2528/30670; GC *rs7041* T=0.4841/5874

Based on the WES results and literature curation, *GC rs4588* was selected for real-time PCR analysis **(Figures 4.5 and 4.6)** performed successfully in 100 patients for genotype-phenotype association analysis in the extended study population. Sanger sequencing was performed on one sample each from the three genotype groups to verify the real-time PCR results. PCR results are shown in **Figure 4.5.** Genotype distribution and allele frequencies of the *GC rs4588* SNV were similar across ethnic groups (data not shown).

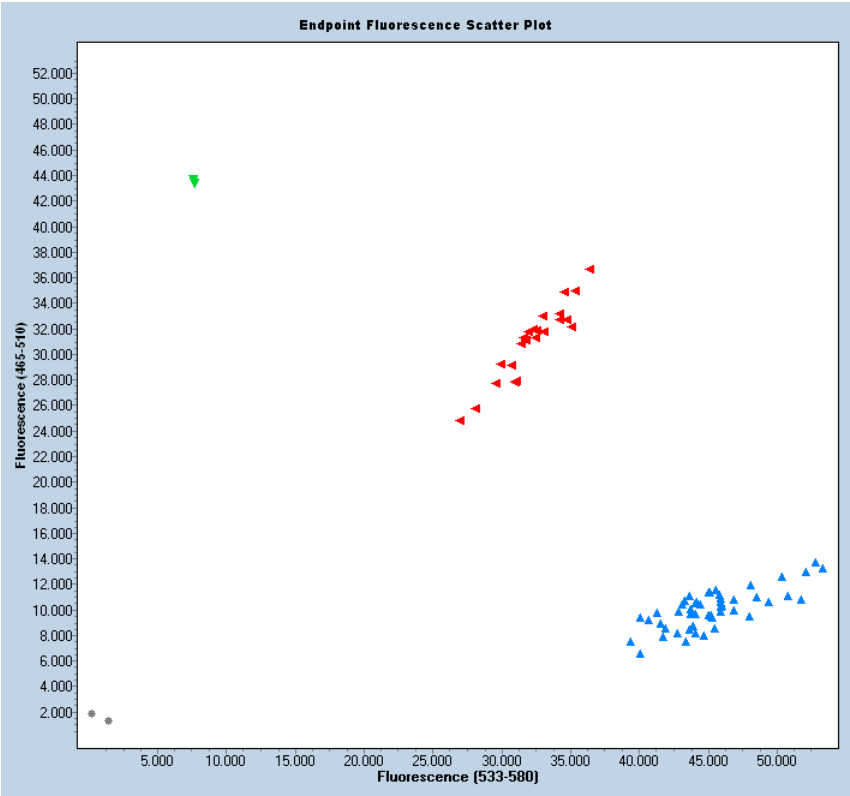


Figure 4.5: Graph showing real time PCR for the GC variant rs4588 using the ABI TaqMan®. [Allele Y (T- FAM labelled) vs Allele X (G- HEX labelled)]. Blue = GG; Red = GT and Green = TT.

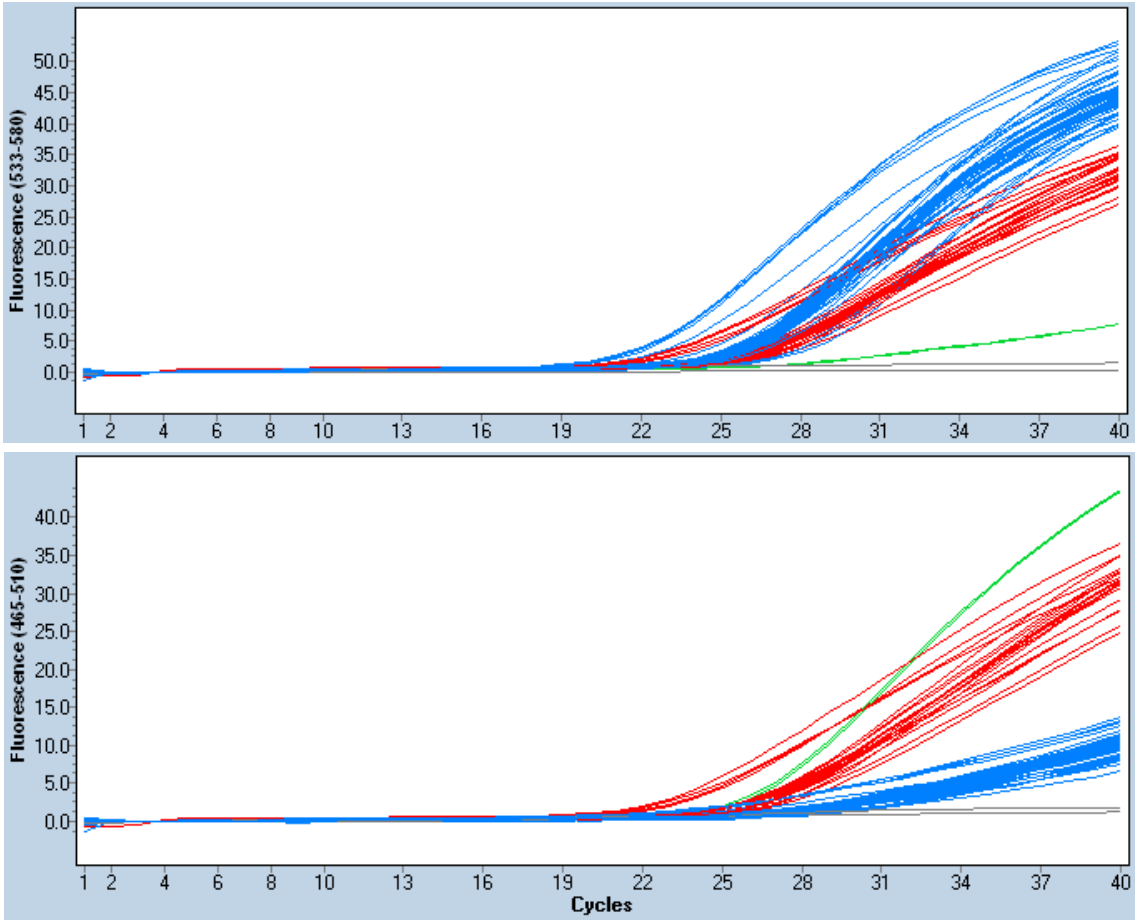
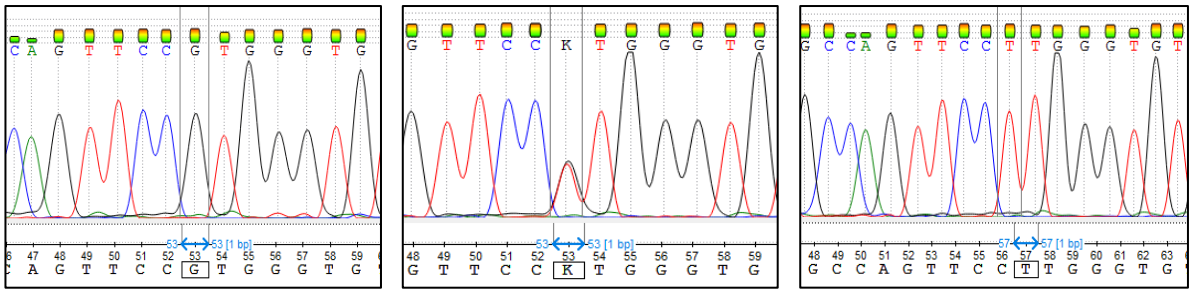


Figure 4.6: Graph showing the fluorescence of real time PCR for the GC variant rs4588 using the ABI TaqMan®. [Allele X (G- HEX labelled) vs Allele Y (T- FAM labelled)], Blue = GG; Red = GT and Green = TT



Genotype GG Genotype GT Genotype TT

Figure 4.7: Confirmation of GC rs4588 by Sanger sequencing

Table 4.3 shows the comparative analysis of clinical characteristics of the study population evaluated for the *GC rs4588* genotypes *GG* vs *GT/TT* combined. The two patients with the homozygous TT genotype had deficient (11.6 ng/mL) or insufficient (21.8 ng/mL) vitamin D levels. No deviation from Hardy-Weinberg equilibrium was observed in the study population (p=0.71).

Table 4.3: Comparison of the clinical characteristics of postmenopausal breast cancer patients with the GC rs4588 GG genotype (n=66) and combined GT (n=32) and TT (n=2) genotypes.

Characteristics	**GC rs4588 GG genotype (n=66)**	**GC rs4588 GT/ TT genotype (n=34)**	***p value***
Vitamin D levels* (n=100)			
Deficient (<20 ng/mL) (n=53)	34 (68)	19 (32)	*0.161*
Insufficient (21-29 ng/ml) (n=43)	31 (72)	12 (28)	
Sufficient (≥ 30 ng/mL) (n=4)	1 (25)	3 (75)	
Body mass Index (n=100)			
Normal (16-25 (kg/m2) (n=24)	19 (79)	5 (21)	*0.145*
Overweight (26-30 kg/m2) (n=22)	17 (77)	5 (23)	
Obese (>30 kg/m2) (n=54)	32 (59)	22 (41)	
Tumour size (n=98)			
<50mm (n=41)	25 (61)	16 (39)	*0.182*
≥50mm (n= 57)	43 (75)	14 (25)	
Tumour type (n=96)			
ICNST (n=71)	49 (69)	22 (31)	*0.879*
ILC (n=13)	8 (62)	5 (38)	
Other (n=12)	8 (68)	4 (32)	
ER status (n=100)			
Positive (n=97)	66 (68)	31 (32)	*1.000*
Negative (n=3)	2 (67)	1 (33)	
PR status (n=100)			
Positive (n= 80)	55 (69)	25 (31)	*0.792*
Negative (n=20)	13 (65)	7 (35)	
HER2 status (n=94)			
Positive (n=12)	10 (83)	2 (17)	*0.326*
Negative (n=82)	54 (66)	28 (34)	
***Bone mineral density (n=53)**			
Normal (n=21)	13 (62)	8 (38)	*0.610*
Osteopenia (<−1.0 > -2.5) (n=18)	10 (56)	8 (44)	
Osteoporosis (≤-2.5) (n=140	32 (60)	21 (40)	

*Data overlap with Baatjes et al. 2019. Percentages are provided in brackets. BMI: body mass index; BMD: bone mineral density; ER: oestrogen receptor status; PR: progesterone receptor status; HER2: human epidermal growth receptor-2

4.3 Analysis of variation on the vitamin D receptor gene

Based on bone mineral density (BMD) classification, 14 patients had osteoporosis, 18 osteopenia and in 21 BMD levels were within the normal range. Using Receiver operator Curve (ROC), a vitamin D level of 17.65 ng/mL was found to best discriminate between patients with normal BMD (-1 > +1) versus patients with osteopenia (< −1.0 > -2.5) or osteoporosis (≤-2.5). The area under the curve (AUC) was only 0.57, indicating poor performance of this cut off **(Figure 4.8).**

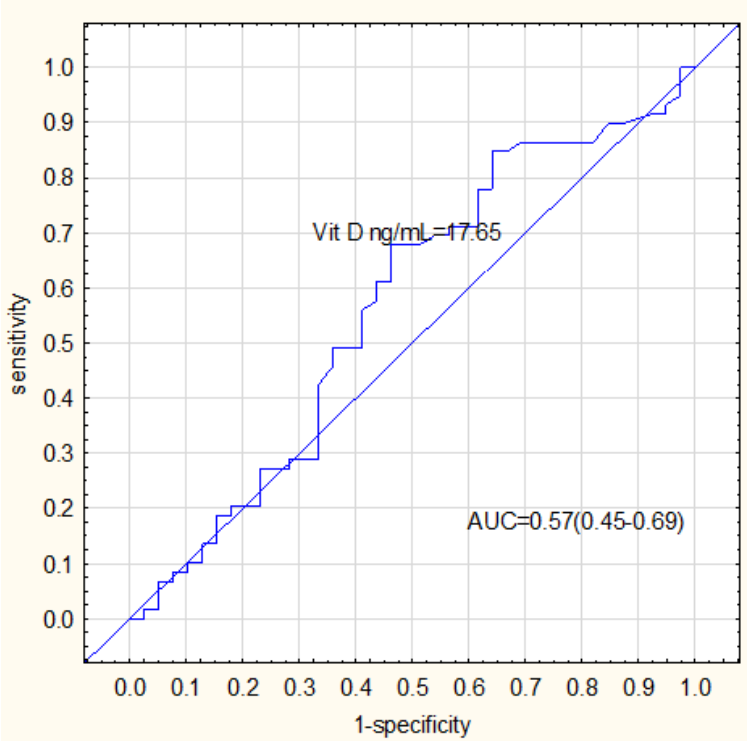


Figure 4.8: Determination of vitamin D concentration that best discriminates between postmenopausal hormone receptor-positive breast cancer patients with normal bone mineral density versus osteopenia/osteoporosis.

Sanger sequencing was performed on DNA samples of all 14 osteoporotic patients for *rs2228570* (FokI) in exon 2, *rs731236* (*TaqI*) in exon 9, *rs7975232* (ApaI) in intron 8 and *rs1544410* (BsmI) in intron 8 in the 3'untranslated region of the *VDR* gene (**Table 4.4)**.

Table 4.4: VDR genotypes for rs2228570, rs731236, rs7975232 and rs1544410 determined by Sanger sequencing in 14 postmenopausal breast cancer patients diagnosed with osteoporosis patients prior to aromatase inhibitor treatment

Samples	**010**	**011**	**014**	**018**	**019**	**044**	**048**	**061**	**063**	**064**	**070**	**079**	**089**	**099**
rs7975232 (ApaI)	GT	TT	TT	TT	TT	TT	GT	GC	TT	TT	GT	GT	TT	TT
rs1544410 (BsmI)	AG	GG	GG	AG	AG	GG	GG	AG	AG	AG	GG	AG	AG	GG
rs2228570 (FoKI)	CT	TT	CT	CT	CT	TT	CT	CT	TT	TT	CT	TT	CT	TT
rs731236 (TaqI)	CC	CT	CT	CC	CC	CC	CC	CT	CT	CT	CT	CT	CT	CC

The exonic *VDR rs731236* (TaqI) and *rs2228570* (FokI) that can be detected by WES covering the coding region of the human genome, was analysed in 55 patients stratified by vitamin D levels. The genotype distribution for these two SNVs shown in **Figure 4.9** (*rs731236, p=0.002)* and **Figure 4.10** (*rs2228570, p=0.001)* differed significantly between patients with osteoporosis, osteopenia and those with normal BMD. For *VDR rs731236* (TaqI) the risk allele is C, found to be absent in breast cancer patients with osteoporosis. The risk allele for *rs2228570* (FokI) was T, with both the minor C and T *VDR* alleles detected at a frequency of 0.42 in the study population.

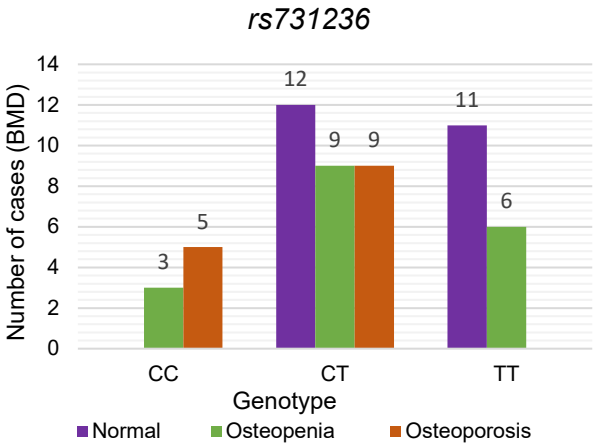


Figure 4.9: Genotype distribution of VDR rs731236 (TaqI) in relation to bone health in the study population of postmenopausal breast cancer patients

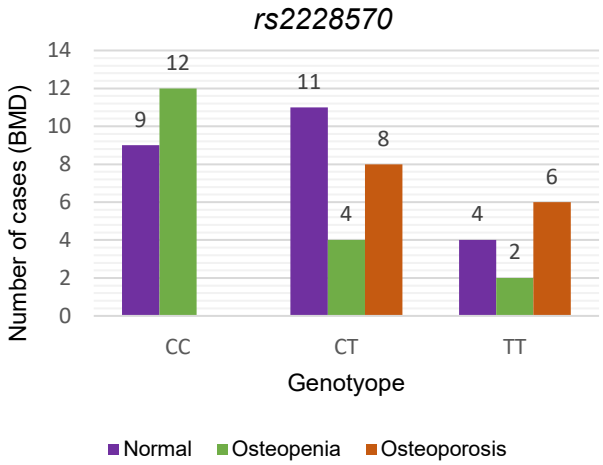


Figure 4.10: Genotype distribution of VDR rs2228570 (FokI) in relation to bone health in the study population of postmenopausal breast cancer patients

The clinical characteristics of postmenopausal breast cancer patients with the *VDR rs731236* (TaqI) genotypes CC (n=8), CT (n=28) and TT (n=17) and *rs2228570* (FokI) genotypes CC (n=20), CT (n=23) and TT (n=12) are shown in tables 4.5 and 4.6, respectively. The risk-associated *VDR rs731236* C-allele and *rs2228570* T-allele were present in all 14 breast cancer patients with osteoporosis ($p<0.01$), while none of the other clinical characteristics were associated with these SNVs.

Table 4.5: Comparison of the clinical characteristics among postmenopausal breast cancer patients genotyped for VDR rs731236 (TaqI)

Characteristics	CC (n=8)	CT (n=30)	TT (n=17)	*p value*
Vitamin D levels* (n=55)				
Deficient (<20ng/mL) n=26	4 (15)	14 (54)	8 (31)	*0.714*
Insufficient (21-29ng/ml) n=19	4 (21)	8 (53)	5 (26)	
Sufficient (≥30ng/mL) n=10		6 (60)	4 (40)	
Body mass Index (n=55)				
Normal (16-25) n=18	4 (22)	8 (44)	6 (34)	*0.639*
Overweight (26-30) n=8		6 (75)	2 (25)	
Obese (>30) n=29	4 (14)	16 (55)	9 (31)	
Tumour size (n=54)				
<50mm n=43	4 (10)	24 (56)	15 (35)	*0.246*
≥50mm n=11	3 (27)	6 (54)	2 (19)	
Tumour type (n=53)				
ICNST n=32	3 (9)	17 (53)	12 (38)	*0.486*
ILC n=13	3 (23)	6 (46)	4 (31)	
Other n=8	1 (13)	6 (74)	1 (13)	
ER status (n=55)				
Positive n=55	8 (15)	30 (54)	17 (31)	
PR status (n=55)				
Positive n= 43	6 (14)	23 (53)	14 (33)	*0.910*
Negative n=12	2 (17)	7 (58)	3 (25)	
HER status (n=52)				
Positive n=3		3 (100)		*0.685*
Negative n=49	6 (12)	27 (55)	16 (33)	
Bone mineral density (n=53)				
Normal (n=23)		12 (52)	11 (48)	*0.002*
Osteopenia (< −1.0 >-2.5) (n=18)	3 (17)	9 (50)	6 (33)	
Osteoporosis (≤-2.5) (n=14)	5 (36)	9 (64)		

Table 4.6: Comparison of the clinical characteristics among postmenopausal breast cancer patients genotyped for VDR rs2228570 (FokI).

Characteristics	CC (n=20)	CT (n=23)	TT (n=12)	*p value*
Vitamin D levels* (n=55)				
Deficient (<20 ng/mL) n=26	9 (35)	11 (42)	6 (23)	
Insufficient (21-29 ng/ml) n=19	9 (47)	6 (32)	4 (21)	*0.639*
Sufficient (≥30 ng/mL) n=10	2 (20)	6 (60)	2 (20)	
Body mass Index (n=55)				
Mean levels (kg/m^2)				
Normal (16-25) n=18	7 (39)	6 (33)	5 (28)	
Overweight (26-30) n=8	4 (50)	3 (38)	1 (12)	*0.774*
Obese (>30) n=29	9 (31)	14 (48)	6 (21)	
Tumour size (n=54)				
<50mm n=43	16 (37)	20 (47)	7 (16)	*0.279*
≥50mm n=11	4 (36)	3 (28)	4 (36)	
Tumour type (n=53)				
ICNST n=32	15 (47)	9 (28)	8 (25)	
ILC n=13	2 (15)	9 (70)	2 (15)	*0.135*
Other n=8	3 (37)	4 (50)	1 (13)	
ER status (n=55)				
Positive n=55	20 (36)	23 (42)	12 (22)	
PR status (n=55)				
Positive n= 43	15 (35)	18 (42)	10 (23)	*0.918*
Negative n=12	5 (42)	5 (42)	2 (16)	
HER status (n=52)				
Positive n=3	1(33)	1 (33)	1 (33)	*1.000*
Negative n=49	18 (37)	21 (43)	10 (20)	
Bone mineral density (n=55)				
Normal (n=23)	8 (35)	11 (48)	4 (17)	
Osteopenia (<−1.0 > -2.5) (n=18)	12 (67)	4 (22)	2 (11)	*0.001*
Osteoporosis (≤-2.5) (n=14)		8 (57)	6 (43)	

4.4 Sequencing of the *CDH1* gene associated with invasive lobular carcinoma

Based on information retrieved from the database, 14 patients had ILC while 79 patients were diagnosed with ICNST. DNA samples of 7 of the 14 patients with ILC were screened using WES for comparison with 6 DNA samples of patients with ICNST. WES was performed using the Ion Torrent platform followed by variant calling including 28 moderate- to high-penetrance breast cancer susceptibility genes, followed by prioritization of the *CDH1* gene for mutation screening. WES results for this gene were compared with that obtained on the Oxford MinION nanopore platform. The results of the *CDH1* gene screening are shown in **Table 4.7**.

Table 4.7: Clinical characteristics and vitamin D-related gene variants identified among postmenopausal breast cancer patients screened for CDH1 gene variants using WES (Ion Torrent) and/or nanopore long-range sequencing (Minion).

SAMPLE	AGE	VITAMIN D (ng/mL)	TUMOUR PATHOLOGY	SNV ID	NUCLEOTIDE	AMINO ACID	MAF (ExAC)	*GC rs4588*	*VDR rs731236 (TaqI)*	*VDR rs2228570 (FokI)*
*007	60	10.3	Lobular	rs1801552	c.T2076C	Ala 692 Ala	T=0.3452	GT	CC	CT
*##019	67	17.4	Lobular	rs33964119	c.C2253T	Asn 751 Asn	T=0.03955	GT	CT	CT
025	68	26.5	Lobular	rs1801552	c.T2076C	Ala 692 Ala	T=0.3452	GT	CT	CC
028	57	23.5	Lobular	rs1801552	c.T2076C	Ala 692 Ala	T=0.3452	GG	CT	CT
*030	54	19.4	Lobular	rs1801552 rs138493551 rs33964119	c.T2076C c.G2205C c.C2253T	Ala 692 Ala Ala 735 Ala Asn 751 Asn	T=0.3452 C=0.0000 T=0.03955	GT	CT	CT
*#038	67	23.1	Lobular	rs376097289	c.A1298G	Asp 433 Gly	G=0.00007	GG	CT	CC
*072			Lobular	rs33964119	c.C2253T	Asn 751 Asn	T=0.03955	GG	TT	CC
*097	63	12.8	Lobular	rs1801552	c.T2076C	Ala 692 Ala	T=0.3452	GT	CT	CT
##105	61	21.6	Lobular	rs1801552	c.T2076C	Ala 692 Ala	T=0.3452	GG	CT	CT
*027	58	28.6	ICNST	rs201511530 rs33969373 rs2229044	c.G671A c.C1896T c.C2634T	Arg 224 His His 632 His Gly 878 Gly	A=0.00003 T=0.01142 T=0.00963	GG	CT	CC
051	56	12.8	ICNST	rs138493551 rs33964119	c.G2205C c.C2253T	Ala 735 Ala Asn 751 Asn	C=0.0000 T=0.03955	GT	TT	CC
069	62	26.7	ICNST	-				GG	TT	CC
078	56	32.2	ICNST	-				GG	CT	TT

*DNA samples screened for pathogenic mutations using both WES and long-range nanopore sequencing using the MinION device; #Pathogenic *BRCA1* mutation-positive *BRCA1* (c.66dupA (p.E23fs); ##Pathogenic *BRCA2* mutation-positive *BRCA2* (c.3881T>A (L1294*); The *BRCA2* mutation detected in Case 019 was previously described by Baatjes (2018) and confirmed at the National Health Laboratory Service, Bloemfontein, prior to genetic counselling and return of results using the PSGT platform.

None of the patients with *CDH1* variants had sufficient vitamin D levels (≥30 ng/mL) **(Table 4.7)**. There was no difference in the number of variants detected between patients with ILC and ICNST. Two missense variants were detected in exon 5 (c.G671A, p.R224H) and exon 9 (c.A1298G, p.D433G) in patients diagnosed with ILC and ICNST, respectively. The result of Sanger sequencing performed for confirmation of the *CDH1* missense variant in case 027 is shown in **Figure 4.11.** WES detected a pathogenic *BRCA1* variant (c.66dupA, p.E23fs) in case 038. This result was confirmed by Sanger sequencing as shown in **Figure 4.12.**

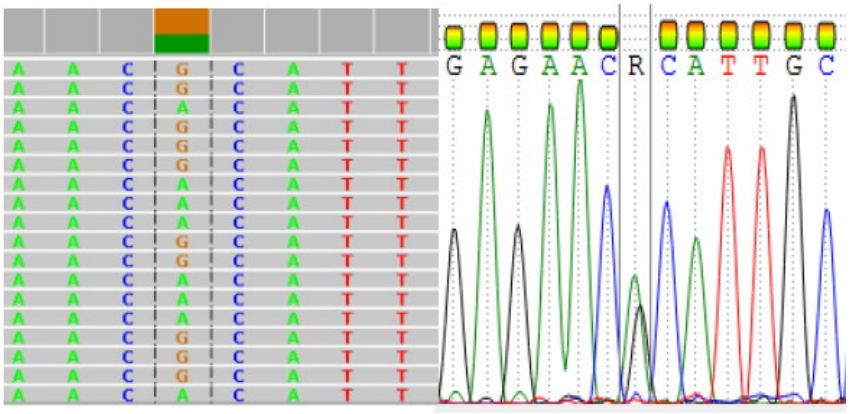


Figure 4.11: Detection of CDH1 c.G671A (p.R224H) in case 027 using WES (left) and confirmed by Sanger sequencing (right).

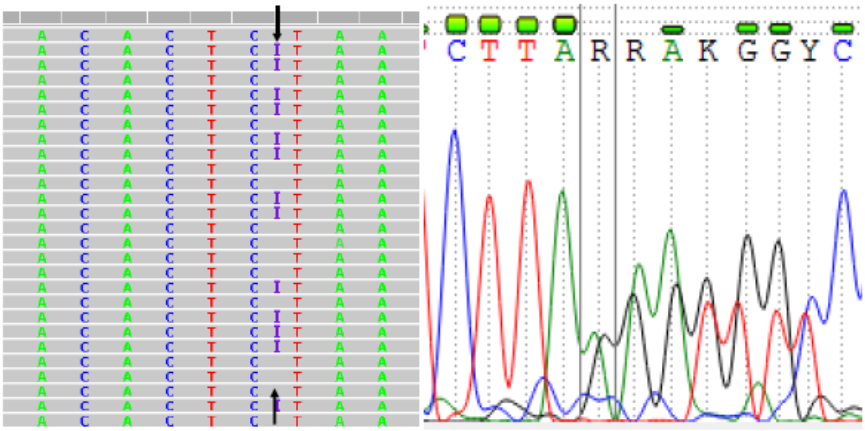


Figure 4.12: Detection of BRCA1 c.66dupA (p.E23fs) in case 038 using WES (left) and confirmed by Sanger sequencing (right).

Results obtained for the *CDH1* gene using WES on the Ion Torrent were compared with the reads generated using nanopore sequencing on the MinION device, as shown with the Integrative Genomics Viewer (IGV) in Figures 4.11 (case 027) and 4.12 (case 038). Figure 4.13 highlights the advantage of using Nanopore sequencing vs. WES for variant discovery, as MinION reads cover the entire genomic sequence of the target gene or *CDH1* gene, whereas WES includes only the exonic / coding regions.

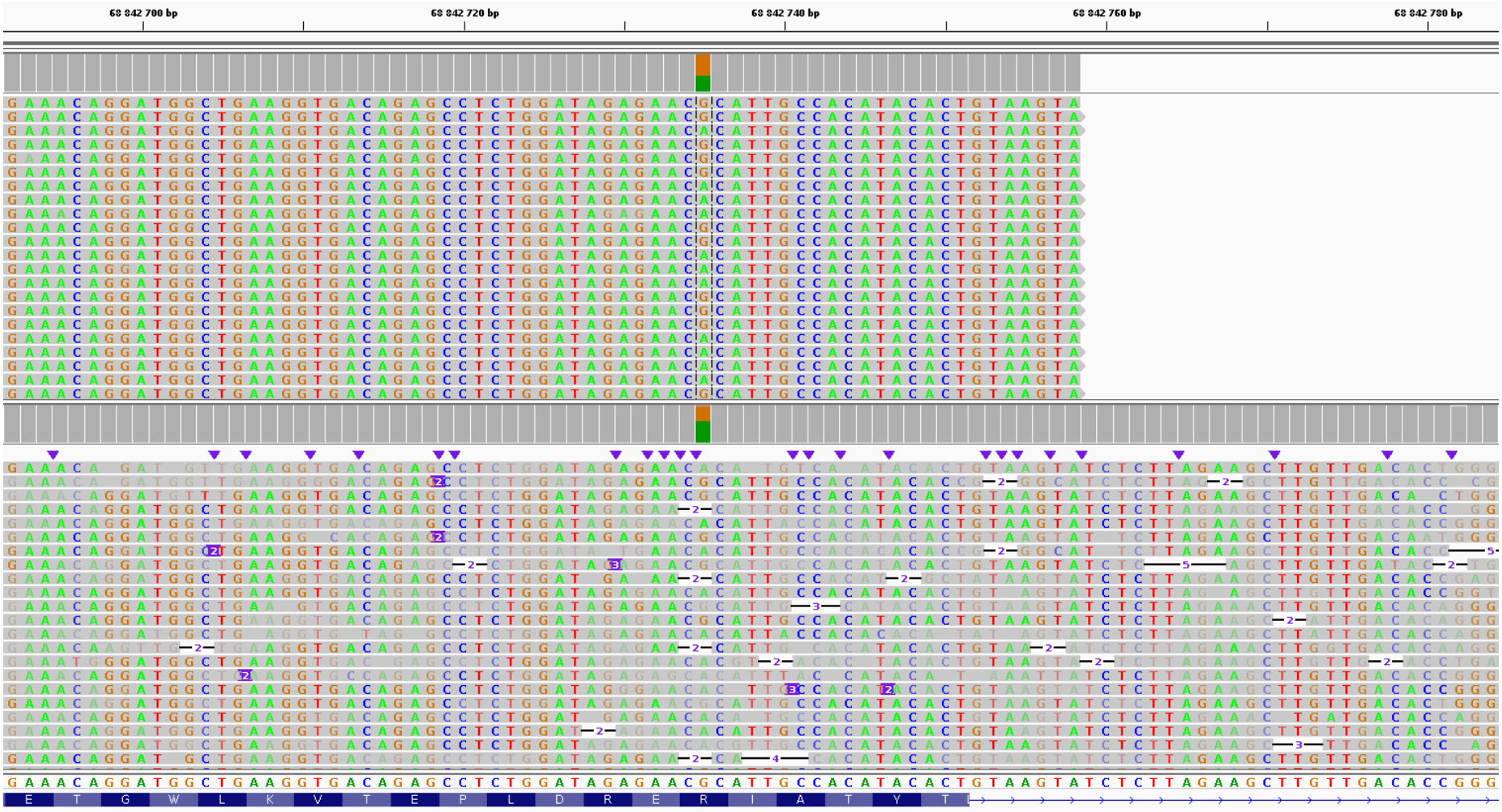


Figure 4.13: Comparative analysis of the CDH1 gene sequenced in Case 027 using WES on the Ion Torrent (top) and long-range nanopore sequencing on the MinION device (below) as viewed using the Integrated Genome Viewer. Comparison of rs201511530 (position chr16:68,842,735) with WES vs Nanopore. Coverage at variant position: WES = 58x vs Nanopore = 64x.

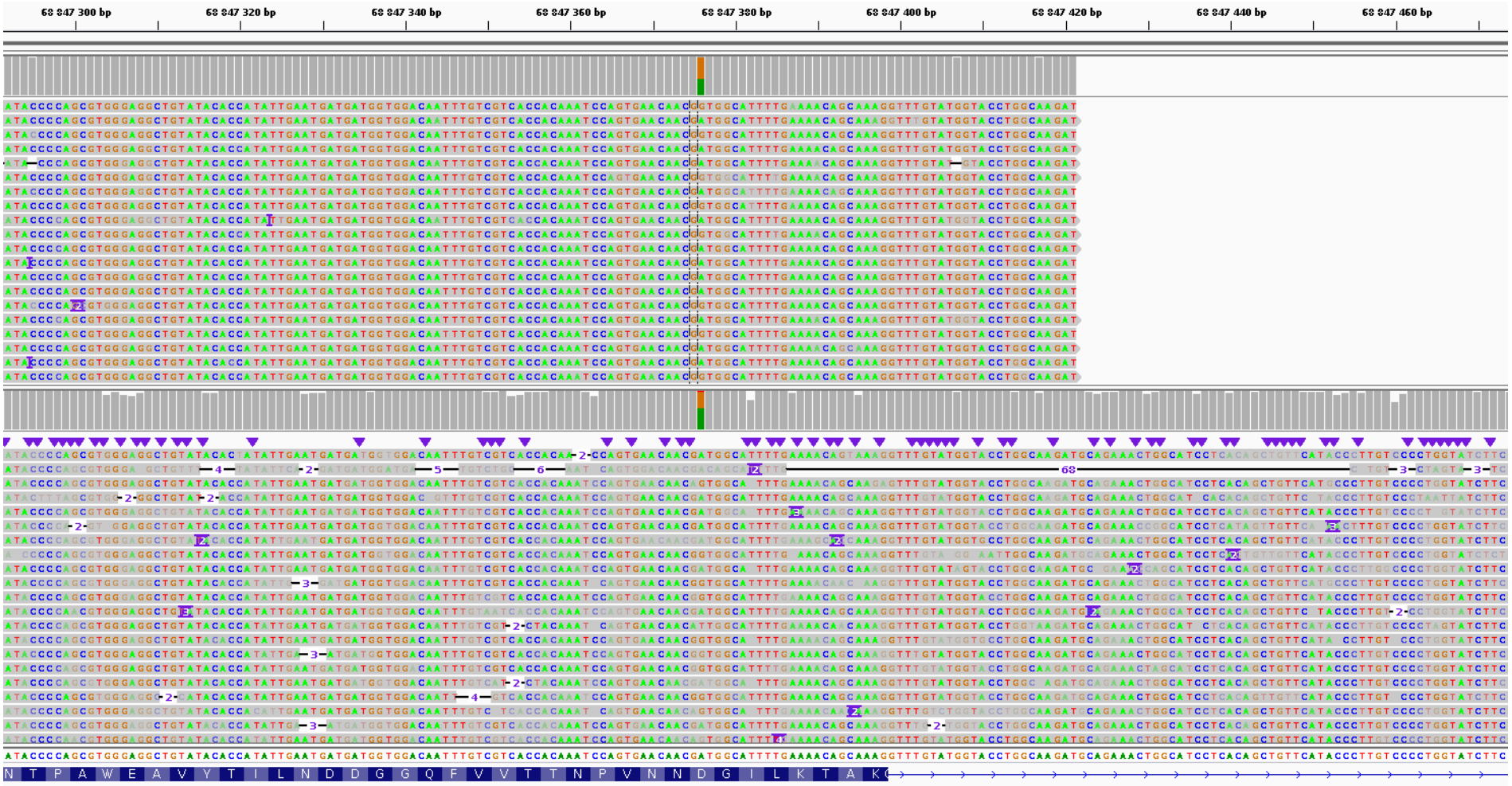


Figure 4.14: Comparative analysis of the CDH1 gene sequenced in Case 038 using WES on the Ion Torrent (top) and long-range nanopore sequencing on the MinION device (below) as viewed using the Integrated Genome Viewer. Comparison of rs376097289 (position chr16:68,847,376) with WES vs Nanopore. Coverage at variant position: WES = 251x vs Nanopore = 86x.

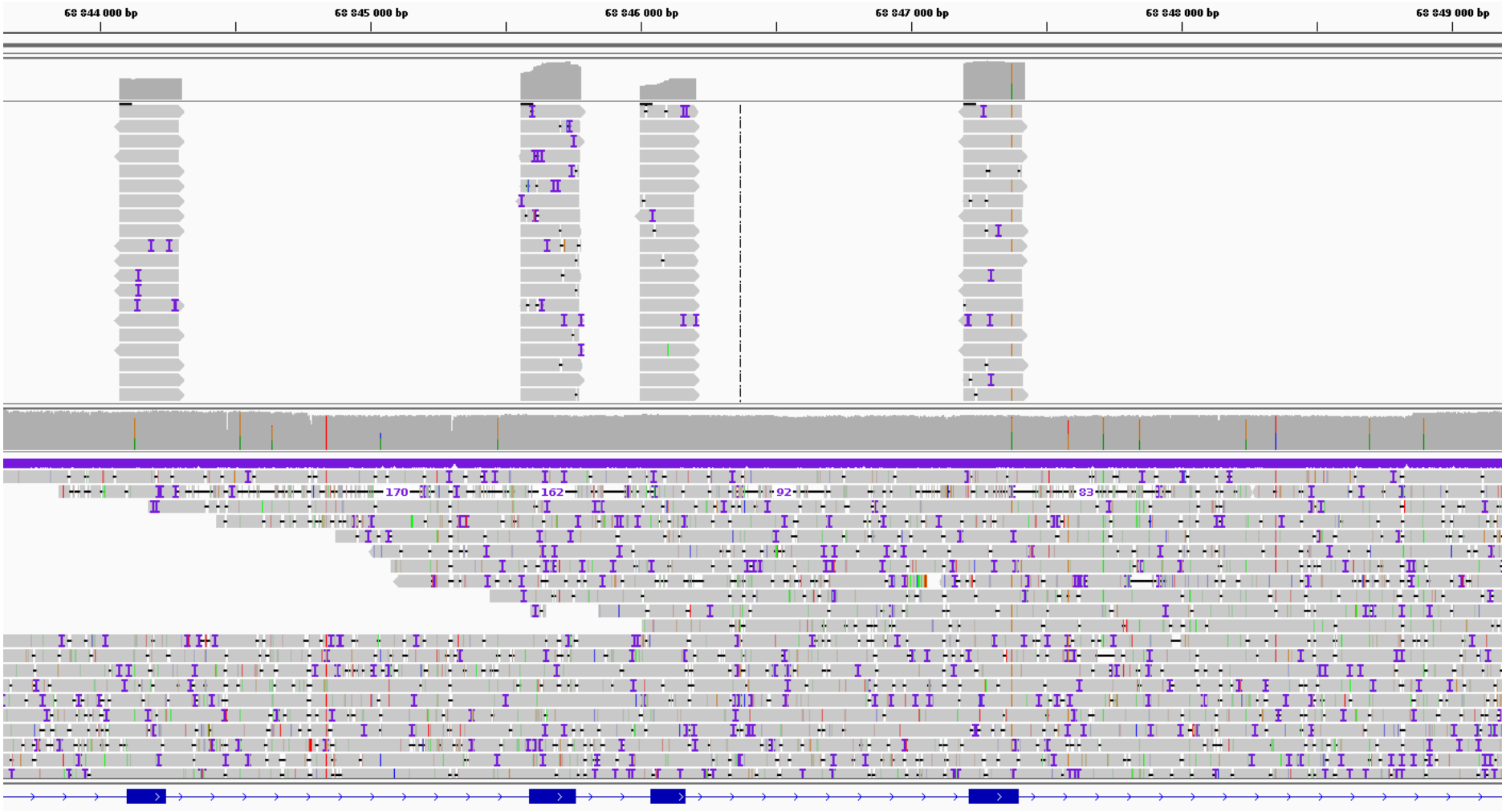


Figure 4.15: Coverage of the CDH1 gene sequenced with WES (top) vs MinION nanopore sequencing (below). Several intronic variants can be seen in the nanopore reads which are missed with WES. (Legend: blue rectangles below nanopore reads = CDH1 exon)

CHAPTER 5

DISCUSSION

The use of an extremes of outcome approach demonstrated that both environmental and genetic risk factors contribute to vitamin D levels, which in turn could affect expression of VDR required for activation of cancer susceptibility genes such as *BRCA1/2* and *CDH1*. This investigation used the database developed during the study of Baatjes et al. (2019), who recently reported that approximately one-third of hormone receptor-positive postmenopausal South African breast cancer patients treated at Tygerberg Academic Hospital between 2014 and 2017, fulfilled international criteria for bone protection during aromatase inhibitor therapy. Only 7% had sufficient vitamin D levels, with compensatory secondary hyperparathyroidism observed in nearly half of the 14 breast cancer patients diagnosed with osteoporosis at baseline.

Determination of BMD furthermore highlighted a significant association between bone fragility and low/normal body weight, with 85% of patients found to be overweight or obese. Follow-up studies after 12 months in a subset of 72 South African patients selected for aromatase inhibitor therapy identified 10 patients with more than 5% bone loss, despite all being obese (Baatjes, 2018). Bone loss was significantly associated with variation in the *CYP19A1* (rs10046) gene, screened for clinically relevant SNVs using WES and literature curation (Baatjes et al., 2017). A similar PSGT approach was used in this study, to firstly identify modifiable clinical and environmental factors contributing to vitamin D levels in the target population of postmenopausal hormone receptor-positive breast cancer patients. Secondly, genetic testing was performed within this context to help distinguish between lifestyle-triggered and inherited breast cancer. Thirdly, therapy-associated co-morbidities were taken into account to determine the most optimal treatment strategy for each patient. There is an urgent need for such comprehensive risk assessment before initiation of aromatase inhibitor

treatment to prevent bone fractures, reported to occur in 18-20% of aromatase inhibitor-treated breast cancer patients after 5-years of follow-up (Edwards et al., 2011). Identification of risk factors for bone loss before the start of aromatase inhibitor treatment will facilitate implementation of preventative measures for improved survival free of medication side effects.

5.1 Clinical assessment of risk factors for vitamin D deficiency

The statistically significant association detected between BMI and vitamin D levels confirmed the effect of excess body weight on vitamin D deficiency, independent of seasonal changes which was shown to affect vitamin D levels in our study population. Inadequate outdoor physical activity noted in the majority of study participants is considered an important risk factor for weight gain and vitamin D deficiency. Of the 116 postmenopausal hormone receptor-positive breast cancer patients included in this study, 59 (53%) were obese and 91 (82%) reported low levels of physical activity. This is a cause for concern, since exercise has been found to be important for release of biologically useful vitamin D trapped in adipose tissue (Hengist et al., 2019). The favourable biochemical changes associated with regular exercise have a protective effect on breast cancer risk and other medical consequences of obesity (Thomas et al., 2017; World Cancer Research Fund/American Institute for Cancer Research, 2018). While there is no evidence that nutritional supplementation promotes weight loss or reduces cancer risk when vitamin D is increased to normal levels, several studies support an association with tumour aggressiveness and clinical outcome.

Abulkhair et al., (2016) reported an increased risk of triple-negative breast cancer in Saudi Arabian breast cancer patients with vitamin D levels equal or below 25 nmol/L, compared to patients with higher levels as determined by chemiluminescence immunoassay. In a cross-sectional study performed in 105 newly diagnosed premenopausal and 95 postmenopausal breast cancer patients, low vitamin D levels were reported in more than 90% of patients, with the highest levels noted in postmenopausal patients (Thanasitthichai, et al., 2015). A cut-off level of <32 ng/mL using high-performance liquid chromatography was used in this study

performed in Thailand, which differed from the standard threshold of <20 ng/mL used in our study population found to be vitamin D deficient in 53% of cases. As vitamin D levels of 40-60 ng/mL have been suggested to be the optimal levels for cancer prevention (Garland, et al., 2009), it is of great concern that only 2 (2%) of the patients in our study population, aged 53 and 58 years, both with normal BMI, had such optimal levels of 42 ng/mL and 43 ng/mL, respectively.

Variation in vitamin D levels due to seasonal changes confirmed that the results of the chemiluminescence immunoassay used in our study was a true reflection of the combined effect of genetic and environmental factors known to affect vitamin D levels. Similar to our results, a study performed in Greece that measured vitamin D levels at different time points over a period of 2.5 years in female patients with osteoporosis, showed the lowest vitamin D levels in winter and the highest levels during summer (Georgios et al., 2015). Since the highest levels of vitamin D were observed during summer, it is clear that latitude is an important factor influencing vitamin D levels (Elizondo-Montemayor et al., 2017). Older women may be vitamin D deficient irrespective of the season as a result of differences in the manner they dress, compared to younger women (Vallejo et al., 2018). Sunlight exposure and UVB rays are responsible for the production of pre-vitamin D in the skin.

Breast cancer is a heterogeneous disease with different treatment requirements based on ER, PR and HER2 status (Turashvili & Brogi, 2017). In a study performed in Egypt including 20 premenopausal and 30 postmenopausal breast cancer patients, significant associations were reported between vitamin D deficiency and tumour size ($p < 0.001$), grade ($p = 0.014$), stage ($p = 0.001$) and HER2 status ($p = 0.002$) (Ismail et al., 2018). However, another research group in Egypt failed to show any association between vitamin D levels and tumour grade, type or HER2 status in 168 (74 premenopausal and 94 postmenopausal) breast cancer patients (El-Shorbagy, 2017). The differences in results may be due to small sample sizes and different methods used for vitamin D analysis, namely Hitachi ROCHE Cobas (electrochemiluminescence immunoassay) and ELISA. A large Belgian study of 1,800 newly

diagnosed breast cancer patients including 1,113 postmenopausal women confirmed the association between vitamin D levels and tumour size, with no evidence for an effect on tumour grade or HER2 status (Hatse et al., 2012). The strengths of the Belgian study were a large sample size and the use of a radioimmunoassay technique considered most accurate for vitamin D analysis. No statistically significant association between vitamin D and tumour characteristics were detected in our study population.

The reasons for discrepancies reported between and within countries have been ascribed to different assay methods, the inability to decide on a universally accepted cut-off level for vitamin D deficiency, differences in reference ranges of instruments used to measure vitamin D levels, diverse patient selection criteria, and heterogeneity of breast cancer that may differ between populations due to diverse genetic backgrounds. Confounding factors such as climate and diet have furthermore been highlighted as reasons for discrepancies in two bone health studies performed in Saudi Arabia. Alkhenizan et al., (2017) could not detect any relationship between vitamin D levels and BMD among 271 males and 1452 females, which contrasted with the results of an earlier study consisting mainly of 448 postmenopausal women from Saudi Arabia (Ardawi et al., 2010). No association was detected between vitamin D levels and BMD in our study cohort.

The risk of osteoporosis linked to vitamin D deficiency is related to a negative calcium balance and increased bone resorption due to a compensatory effect associated with secondary hyperparathyroidism (Prieto-Alhambra et al., 2011; Servitja et al., 2012). Lack of association with BMD subcategories in South African breast cancer patients first noted by Baatjes et al. (2019), was ascribed to the universally low vitamin D levels. The same direction of action on bone health imposed by high/sufficient vitamin D levels and high BMI, which in turn correlates with low vitamin D known to be a risk factor for osteoporosis, should furthermore be considered in this context. Complex medical conditions such as breast cancer and osteoporosis are characterized by multiple intermediate phenotypes involved in their pathogenesis (Blanco-

Gómez et al., 2016), with age, gender and genetic risk factors identified as the most important non-modifiable risk factors. The impact of modifiable risk factors on bone health needs to be minimised by optimal vitamin D levels, a healthy body weight, regular outdoor activities, adequate calcium intake, avoidance of both active and passive smoking and excessive alcohol consumption. Global consensus recommends bone protective therapy in all women with a BMD T-score of less than -2 SD; or with a T-score between -1.5 and -2 SD in the presence of additional risk factors related to the above-mentioned modifiable risk factors, including age above 65 years, a personal history of fragility, a family history of hip fracture and a course of steroid therapy for longer than three months.

Although our findings are limited to the results obtained with single measurement of vitamin D levels available in the research database, it is considered adequate (Jorde et al., 2010) for the first phase of this study. We focused on the determination of modifiable clinical and environmental risk factors for vitamin D deficiency that may, in turn, have a negative effect on genetic variation affecting the vitamin D pathway or expression of the VDR (Pervin et al., 2013).

5.2 Whole exome sequencing for identification of vitamin D-lowering genetic variants

A pilot WES study was performed using DNA samples from 10 patients with extremes of vitamin D levels (≤12 ng/mL and >30 ng/mL) to identify SNVs associated with low vitamin D levels in our study cohort of postmenopausal hormone receptor-positive breast cancer patients. The bioinformatics tools used to facilitate variant detection in vitamin D pathway genes, including *GC, DHCR7, CYP2R1, CYP27B1* and *CYP24A1,* revealed two missense variants in the *GC* gene (*rs4588* and *rs7041)*. The risk-associated T-allele of *GC rs4588* was detected in all five patients with very low vitamin D levels (≤ 12 ng/mL), including one TT homozygote and two compound heterozygotes with both *rs4588* and *rs7041*. Two of the five breast cancer patients with vitamin D levels in the highest extreme of normal values (>30

ng/mL) were heterozygous for the *GC rs4588* genotype, and three were heterozygous for *GC rs7041.* None of these patients with high vitamin D levels were compound heterozygotes.

Notably, none of the patients with vitamin D levels above 30 ng/mL were obese; one patient was classified as overweight based on BMI. In contrast, four of the five patients with ultra-low vitamin D levels were obese; one patient had a normal BMI. The effect of seasonal variation as confirmed in the target patient group was also evident, since levels measured in winter were only reported in two patients in the low vitamin D group. The power of PSGT incorporating WES to help distinguish between causal genetic and environmental factors, or a combination of both, was evident in at least one patient (case 004). In this severely vitamin D deficient patient with the lowest vitamin D level of all study participants measured in the winter (4.82 ng/mL), BMI was classified as obese and compound heterozygosity detected for *GC rs4588* and *rs7041.* The ultra-low vitamin D levels detected in this case therefore appears to be the result of a combined effect of BMI, reduced sun exposure during winter and vitamin D lowering SNVs.

GC rs4588 was selected for real-time PCR analysis in the extended study population based on the literature study and above-mentioned WES results interpreted in a clinical context. Of the 100 patients successfully genotyped for this SNV, 66% were homozygous for the G-allele, 32% heterozygous and 2% homozygous for the risk-associated T-allele. Since no statistically significant association was detected between this *GC* allele and vitamin D levels, the importance of modifiable environmental factors such as high BMI and inadequate outdoor physical activity (that may mask the gene effect as confounders) were confirmed in the target population of postmenopausal hormone receptor-positive breast cancer patients.

The same PSGT approach was previously used by Baatjes et al., (2017) for selection of *CYP19A1 rs10046* following WES performed in an initial 5 patients with low vitamin D levels. Proof of concept was subsequently provided in a subset of 72 patients (included in our study population) followed-up after 12 months of aromatase inhibitor therapy: *CYP19A1 rs10046*

AA-homozygotes were found to be 7-10 times more likely to have a higher percentage bone loss at the total hip (CI of 1.10- 49.34, p=0.04) or the lumbar spine (CI of 1.77- 65.83, p=0.01) compared to patients without this allele (Baatjes, 2018). The clinical relevance of *CYP19A1 rs10046* homozygosity was confirmed by absence of this genotype in DNA samples of 34 breast cancer patients without bone loss at the lumbar spine, after one year on aromatase inhibitors. High WES coverage allowed accurate detection of *CYP19A1 rs10046* within a comprehensive pharmacogenetics and diagnostic context, as one of the 10 patients who experienced more than 5% aromatase inhibitor-induced bone loss also tested positive for a pathogenic *BRCA2* mutation (c.582G>A). This patient was homozygous for the A-allele of the functional *CYP19A1 rs10046* polymorphism, which support the use of comprehensive DNA sequencing methods such as WES providing a resource for both variant discovery and differential diagnosis in a clinical context.

A major limitation of many studies involving vitamin D pathway genes is that vitamin D levels are not measured as an intermediate phenotype between gene and disease. This is required for appropriate clinical interpretation of functional effects associated with low-penetrance SNVs such as *GC rs4588,* occurring at a relatively high frequency in the general population. Interdisciplinary studies enabled by PSGT or other platforms that aim to translate genetics/genomics discoveries into clinical practice is essential to extract useful information from laboratory-based research that could have a positive impact on public health. Stratification by cancer type previously showed that *GC rs4588* was significantly associated with breast cancer risk, while rs7041 was associated with a significantly increased risk of prostate and lung cancer (Zhu et al., 2019). However, in a study of 500 postmenopausal breast cancer patients and matched controls performed in the United States, no statistically significant association was detected between GC *rs4588* and cancer risk (McCullough et al., 2007). Notably, BMI shown to be an important risk factor in differential diagnosis of genetic and lifestyle-related breast cancer in our study was not taken into account as a potential confounding factor. The effect of high calcium intake on reduction of breast cancer risk

mediated by genetic variation in *VDR* that is activated by vitamin D could be demonstrated, which suggests significant gene-diet interaction. Many contradictory findings such as these have been described in the literature that highlight the knowledge gap created by non-integration of genetics with pathology and other health disciplines where necessary (Anderson, et al., 2011; Abbas et al., 2008; Carpenter et al., 2013; Jorde et al., 2015; Maneechay et al., 2015; Tagliabue et al., 2015). The use of different germline DNA extraction (e.g. blood, saliva, cheek swabs) and genotyping methods such as the MassARRAY iPLEX Gold Sequenom Platform used in some of the aforementioned studies differ from the methodology applied in our study population. This is not expected to affect the outcome of comparative studies as all laboratory methods are validated or standardised before implementation and statistical tools used to help identify genotyping errors.

The majority of our female study participants were obese, with an even distribution of the *GC rs4588* GG (n=32) and GT/TT (n=22) genotypes across the three subcategories of obese, overweight and normal/low BMI respectively. Contrasting findings of an association between this SNV and obesity/vitamin D levels (Almesri et al., 2016; Santos et al., 2017) could be explained by differences in exclusion and inclusion criteria used in these studies.

GC rs4588 was not associated with tumour type, ER, PR and HER status in our study. This finding would be expected as our study population was selected by ER status, similar to results reported in a German study for *GC rs4588* (Abbas et al., 2008). Determination of the influence of genetics on vitamin D levels was limited by the relatively small sample size of 100 breast cancer patients. Another limitation was the lack of a control group for comparison of genotype distribution and allele frequencies.

5.3 Analysis of VDR gene variants *rs731236* and *rs2228570*

Genetic risk factors for osteopenia/osteoporosis include genetic variation in the vitamin D pathway and *VDR* genes. Expression of VDR is associated with bone mass, but many studies

failed to show a significant association with vitamin D levels (Kim et al., 2015). In this study the *VDR rs731236* (TaqI) and *rs2228570* (FokI) were analysed in 55 postmenopausal breast cancer patients using WES. The risk-associated allele of *VDR rs731236* (TaqI) is C, while the risk-associated allele of *rs2228570* (FokI) is T (Beysel et al., 2018), both detected at a minor allele frequency of 0.42 in our study population. The genotype distribution of these SNVs differed significantly between South African breast cancer patients with osteoporosis, osteopenia and those with normal BMD included in this study.

None of the postmenopausal breast cancer patients with osteoporosis studied had the *VDR rs731236* TT-genotype associated with methylation status of the *VDR* gene (Andraos et al., 2011), while the risk-associated C-allele was present in all 14 patients. Similarly, none of the osteoporotic patients had the *VDR rs2228570* CC genotype, while the risk-associated T-allele was present in all 14 patients. The functional effect of *VDR rs2228570* (FokI) results in a shorter protein receptor and higher transcriptional activity (Arai et al., 1997) with clinical relevance supported by some (Sinotte et al., 2008; Tang et al., 2009), but not all studies (Rashid et al., 2015; Shahabi et al., 2018). The finding of Andraos et al. (2011) showing differential methylation of *VDR rs731236* according to genotype (TT 0%, TC 50%, CC 100%) may provide an explanation for population differences in disease patterns frequently ascribed to ancestral geographic origins and access to healthcare.

Binding of vitamin D to VDR results in formation of the vitamin D-VDR axis that interacts with DNA to suppress or promote vitamin D target genes. Changes in this axis caused by low levels of vitamin D or (epi)genetic variation in the *VDR* gene, could increase VDR expression in fat cells that in turn reduces energy expenditure (Kazemian et al., 2019). A high BMI was previously found to be protective against bone loss due to release of osteogenic hormones from adipose tissue and increased mechanical loading (Mo et al., 2017). However, as noted by Baatjes et al. (2019), the protective bone effect imposed by obesity was not evident in all 14 of our breast cancer patients with osteoporosis. Seven had a low/normal BMI, four were

overweight and three obese. This may be due to undefined genetic risk factors or the negative effect of inflammatory cytokines produced by excessive fat mass (Kimble et al., 1996). In the first controlled trial investigating the effect of vitamin D supplementation in breast cancer survivors with low plasma vitamin D levels, changes in inflammatory biomarkers were at least in part dependent on variation in the VDR gene (Kazemian, Akbari, et al., 2019). VDR mediates the mechanism of action of vitamin D on genomic stability related to regulation of key genes involved in inflammation, metastasis, and many other processes involved in overall health and carcinogenesis (Narvaez et al., 2014).

Based on the observation of a significant association between breast tumour subtypes and variation in the *VDR* gene, a role of vitamin D in the development of breast cancer was suggested (Shi et al., 2016). Al-Daghri et al. (2017) furthermore demonstrated that *VDR* SNVs have a significant effect on response to vitamin D supplementation in patients with type II diabetes. This finding raised the possibility that *VDR* genotyping could identify patients most likely to benefit from vitamin D treatment. Indeed, patients with the risk-associated *VDR* *rs731236* (TaqI) genotype experienced the most benefit, while individuals with the *rs2228570* (FokI) CC genotype appear to require higher doses of vitamin D to increase their levels to the normal range. These findings support a genotype-targeted approach to vitamin D supplementation, confirming that knowledge of both genetic and pathology test results is required for health optimisation (Al-Daghri et al., 2017).

5.4 Analysis of *CDH1* associated with invasive lobular carcinoma

The *CDH1* gene previously found to be activated by VDR was analysed in patients stratified by tumour type, based on the finding that none of the South African postmenopausal breast cancer patients with ILC had sufficient vitamin D levels. Previous epigenetic studies identified a significant role of *CDH1* hypermethylation in the pathogenesis of breast cancer (Caldeira et al., 2006; Liu et al., 2016). WES of the *CDH1* gene using the Ion Torrent platform was performed using germline DNA of 12 patients, including 9 with ILC and 4 with ICNST. The

CDH1 WES results of 7 patients (6 ILC, 1 ICNST) were compared with the performance of the Oxford Nanopore Technologies MinION platform, showing identical results for the variants studied.

WES identified several common SNVs and two rare *CDH1* variants of uncertain clinical significance (VUS). *CDH1* c.G671A (p.R224H) detected in a breast cancer patient with ICNST was classified as benign. This was based on the fact that pathogenic germline *CDH1* variants are associated with ILC and diffuse gastric cancer, but not ICNST. *CDH1* c.A1298G (p.D433G) was identified in a South African patient with ILC, together with a pathogenic *BRCA1* variant simultaneously detected by WES. Although this finding supports a likely benign classification as reported in the ClinVar database, a family history of stomach cancer reported by the patient raised the possibility of a *CDH1* modifier gene effect on *BRCA1* gene expression in this hormone receptor-positive breast cancer patient.

The finding that *BRCA1* expression is critical for mediating the beneficial biological impact of vitamin D on breast cancer growth (Pickholtz et al., 2014) highlights the role of tumour pathology and the genetic background of individual patients to achieve effective vitamin D-based therapies. Grotsky et al. (2013) reported interaction of vitamin D with the 53BP1 protein to eliminate invasive breast cancer cells that lacks BRCA. Disruption of the transcriptional activation pathway affecting the *BRCA1* promoter leads to development of both breast and prostate cancers (Grotsky et al., 2013). Further studies are warranted to determine whether germline variants in the *CDH1* gene play a role in development of hormone receptor-positive breast cancer in *BRCA1* mutation carriers. Since most breast cancers with pathogenic germline *BRCA1* mutations are triple-negative, further studies were performed to define the clinical characteristics and pathological features of hormone receptor-positive *BRCA1* breast cancers (Lakhani et al., 2010; Tung et al., 2010). It was found that these *BRCA1* carriers are generally older (>50 years) postmenopausal patients and display intermediate tumour pathology that lies between ER-negative *BRCA1* cancers and ER-positive sporadic breast

cancers. This suggested a unique mechanism of cancer development, or an incidental finding considered unlikely in our case, as the same mutation has previously been detected in at least 5 South African breast cancer patients (N van der Merwe, unpublished results).

The pleiotropic nature of *CDH1* linked to different medical conditions (Figueiredo et al., 2019) complicates clinical interpretation of variants identified in this gene (Benusiglio, 2017). This was evident in the case of breast cancer susceptibility discussed at a workshop of the UK Cancer Genetics Group, where a consensus decision was made to exclude *CDH1* from the gene panel of high-risk genes routinely screened in patients with inherited breast/ovarian cancer (Taylor et al., 2018). The reasons for excluding this gene was to avoid the need for reviewing the histopathology reports to confirm an invasive lobular breast cancer (ILC) in patients with a variant of uncertain clinical significance in the *CDH1* gene, as well as the uncertainties encountered in families with no history of ILC or diffuse gastric cancer. It was however agreed that testing for *CDH1* and variant classification should be performed in relevant cases according to existing clinical guidelines. These include amongst other genes *ATM* and *CHEK2*, which both confer a moderately increased risk of breast cancer. The UK Cancer Genetics Group recommended that only truncating variants should be reported in these two genes, in addition to *ATM* c.7217T>G p.(Val2424Gly) that confers an increased risk of breast cancer. Insufficient evidence was found for a significant risk of breast cancer associated with *NBN, BRIP1* and *BARD1*, which were also excluded from the clinical sequencing gene panel used in the UK (Taylor et al., 2018).

WES beyond *BRCA1/2* gene screening has proven its worth in rapid genetic diagnosis and screening for research and clinical service delivery purposes (Singleton, 2011), while long-range nanopore sequencing is a relatively new emerging technology. Considering the portability and affordability of the MinION, it will be a useful sequencing platform in resource-limited settings (Plesivkova et al., 2019), where the cost of high-throughput laboratory-based apparatus is a major limiting factor. The error rate was originally identified as the biggest

problem being faced by nanopore sequencers, but this has gradually been resolved by upgrading the chemistry of the flow cells. Version R6 had an accuracy of only 60% when it was initially launched, whereas with availability of the most recent version R9.5 accuracy has improved to 99% (Plesivkova et al., 2019; Rang, et al., 2018). Streamlining of the library preparation process to become fully automated would result in wider application of the MinION device (Plesivkova et al., 2019). Discontinuation of the EPI2ME cloud-based basecaller and introduction of the MinKNOW integrated software and of the Guppy basecaller are additional measures to reduce the error rate of this platform (Plesivkova et al., 2019).

Nanopore sequencing of LR-PCR amplicons is a robust and sensitive method for variant discovery in the entire genomic region of target genes, limited to the *CDH1* gene in this study. No errors were detected in the *CDH1* gene when the nanopore sequencing results were compared with the WES reads generated in our study, while sequencing beyond the coding regions covered by WES is considered an advantage for future research of variants detected outside the *CDH1* coding region (data not shown). Further studies are warranted to validate nanopore sequencing using the MinION or other devices developed for the same purpose, based on the advantages associated with long reads that can fill the gaps left by WES, which does not cover introns and regulatory regions of genes that may contain clinically relevant gene variants.

Genetic counselling was recommended for the *BRCA1*-mutation-positive patient to determine the appropriateness of cascade testing on both the paternal and maternal sides of the family. This is important to facilitate *CDH1* variant classification, and most importantly, to disclose the detection of a pathogenic *BRCA1* mutation (c.66dupA, p. E23fs) in the index patient using an approved protocol.

5.5 Data Integration using the PSGT platform

The following results obtained in this study add value to the PSGT platform in relation to postmenopausal hormone receptor-positive breast cancer (diagnosis) in patients at risk of osteoporosis (co-morbidity) that may be triggered by treatment with aromatase inhibitors:

1) Lifestyle-triggered vitamin D deficiency: obesity and lack of outdoor physical activities
2) Genetic predisposition for osteoporosis: variation in the *VDR* gene
3) Tumour histopathology and familial risk: Clinical relevance of pathogenic *BRCA1/2* mutations interpreted in relation to VUS identified in the *CDH1* gene
4) Nanopore sequencing using the MinION device confirming the accuracy of long-range sequencing beyond gene coding regions covered by WES

To our knowledge, this is the first study using WES to investigate the significance of genetic variation in vitamin D-related genes in postmenopausal South African breast cancer patients at risk of aromatase inhibitor-induced bone loss. The ability of PSGT to address different aspects of the same disease is a major benefit as demonstrated by WES to simultaneously target familial, lifestyle risk and treatment implications of vitamin D deficiency that has a significant impact on VDR expression. In the absence of vitamin D this receptor remains in the cytoplasm and prevents activation of target genes such as *CDH1* and *BRCA1*. Variation in the *VDR* gene provides a genetic link between vitamin D dysfunction and shared disease mechanisms in chronic diseases of lifestyle, including hormone receptor-positive breast cancer and osteoporosis. These findings strengthen the importance of knowledge relating to both genetic and pathology test results in the attainment of effective therapies.

Strengths and limitations of the study

The value of PSGT in healthcare delivery lies in early detection of patients at increased risk of breast cancer recurrence and drug-induced side effects. Despite a relatively small sample size, the expected seasonal variation in vitamin D levels were observed and confirmed the reliability of the method used to measure vitamin D levels. Furthermore, all breast cancer patients included in this study were screened by a single clinician (KJ Baatjes), providing a

clinically enriched database resource. Lack of follow-up to monitor vitamin D status in relation to bone mineral density limits the interpretation of our results in relation to bone health. Further studies are warranted to increase the number of vitamin D related genes, including pigment genes shown to be associated with vitamin D level, which may also help to exclude population substructure as an important confounder in genetic studies. This study has improved our understanding of gene-environment interactions that became evident at the individual level in this study, as a result of the risk profile defined in the larger target population selected for analysis.

In eligible cases, the PSGT platform was used to integrate findings considered relevant to breast cancer diagnosis with comorbidities and treatment considerations discussed with the treating clinicians and a registered genetic counsellor. Classification of the pathogenic *BRCA1* variant is highly unlikely to change in future, while the *CDH1* VUS detected in the same patient may require recontact of patients in the event of reinterpretation, following family screening or extended testing of additional unrelated breast cancer patients (Bombard et al., 2019).

CHAPTER 6

CONCLUSIONS

Advances in genomic technologies have the potential to improve breast cancer risk management across the continuum of care, ranging from early stage to metastatic disease and survivorship. Lack of an effective data integration system led to development of the PSGT platform (Kotze et al., 2015) applied in this study to facilitate research translation. This concept involves both the germline genome that determines familial risk and governs treatment toxicity, and the somatic genome that determines prognosis and the tumour's response to therapy.

Replacement of tamoxifen with aromatase inhibitors as the first-line treatment of hormone receptor-positive breast cancer in postmenopausal patients has extended patient survival at the cost of increased bone loss. Vitamin D deficiency is a modifiable risk factor for osteoporosis recently identified as a major co-morbidity in postmenopausal hormone receptor-positive breast cancer patients treated at Tygerberg Academic Hospital, South Africa (Baatjes et al., 2019). Since vitamin D deficiency reported in the majority of these patients remains undetected and untreated, this study used data from the same patient cohort to identify determinants of vitamin D levels, which in turn was correlated with tumour characteristics and bone mineral density with the aim to translate research findings into clinical practice. International guidelines are available for early detection or prevention of fractures to enable careful assessment of the benefits and harms to make informed decisions on choice of hormonal therapy.

Previous studies indicated that the *GC rs4588* (c.1307C>A, T436K) is a functional SNV causing an estimated 3.5 ng/mL decrease in vitamin D levels (O'Brien et al., 2018). WES identified this variant as a likely genetic contributing factor to ultra-low vitamin D levels in all 5 South African breast cancer patients, of whom 4 were obese. Extended *GC* genotyping in the total study cohort failed to demonstrate a significant association with vitamin D levels, implying that lifestyle factors such as obesity is the major contributor to vitamin D deficiency

in postmenopausal hormone receptor-positive breast cancer patients. Since vitamin D accumulates in adipose tissue where the majority of the circulating vitamin D is sequestered (Hengist et al., 2019), exercise programs need to be implemented alongside chemoprevention or endocrine treatment to mobilise vitamin D from adipose tissue. Smoking and alcohol consumption (Brouwer-Brolsma et al., 2013; Shi et al., 2014) are also important modifiable lifestyle factors affecting vitamin D levels, which forms part of the information used in the questionnaire-based assessment using the PSGT platform. The beneficial effect of vitamin D on both aromatase inhibitor-induced joint pain and bone loss (Prieto-Alhambra et al., 2011) led to recommendations of a weekly dose of 10,000 IU vitamin D together with a daily dose of 1,000 mg calcium (Rizzoli et al., 2012). Based on the results obtained in this study routine measurement of vitamin D levels in breast cancer patients attending Tygerberg Academic Hospital is important to optimizing bone health during aromatase inhibitor treatment, especially during winter. Decreased vitamin D levels correlated significantly with increased BMI in our study population, independent of seasonal variation.

The realization that variation in skin pigment genes rather than measured skin pigmentation may underly differences in vitamin D levels between African and European populations (Datta et al., 2019) implies that the use of race as a proxy for the root cause of aggressive forms of breast and prostate cancer (Akinyemiju et al., 2017) may be a missed opportunity to optimise public health. Considering the tumour heterogeneity of breast cancer and inconsistent outcomes associated with vitamin D-related SNVs, we propose the need to routinely consider gene-environment interactions for a better understanding to improve treatment strategies with a proposed algorithm shown in Figure 6.1. Without such a well-defined strategy using PSGT to move beyond the limitations of genetics or pathology when applied in isolation, personalised genomic medicine will remain elusive. The application of advanced molecular technologies comes with the responsibility to act on the information they provide and, most importantly, the need for a system to manage patient expectations before embarking on genome-scale analyses (Kotze et al., 2013; van der Merwe et al., 2017).

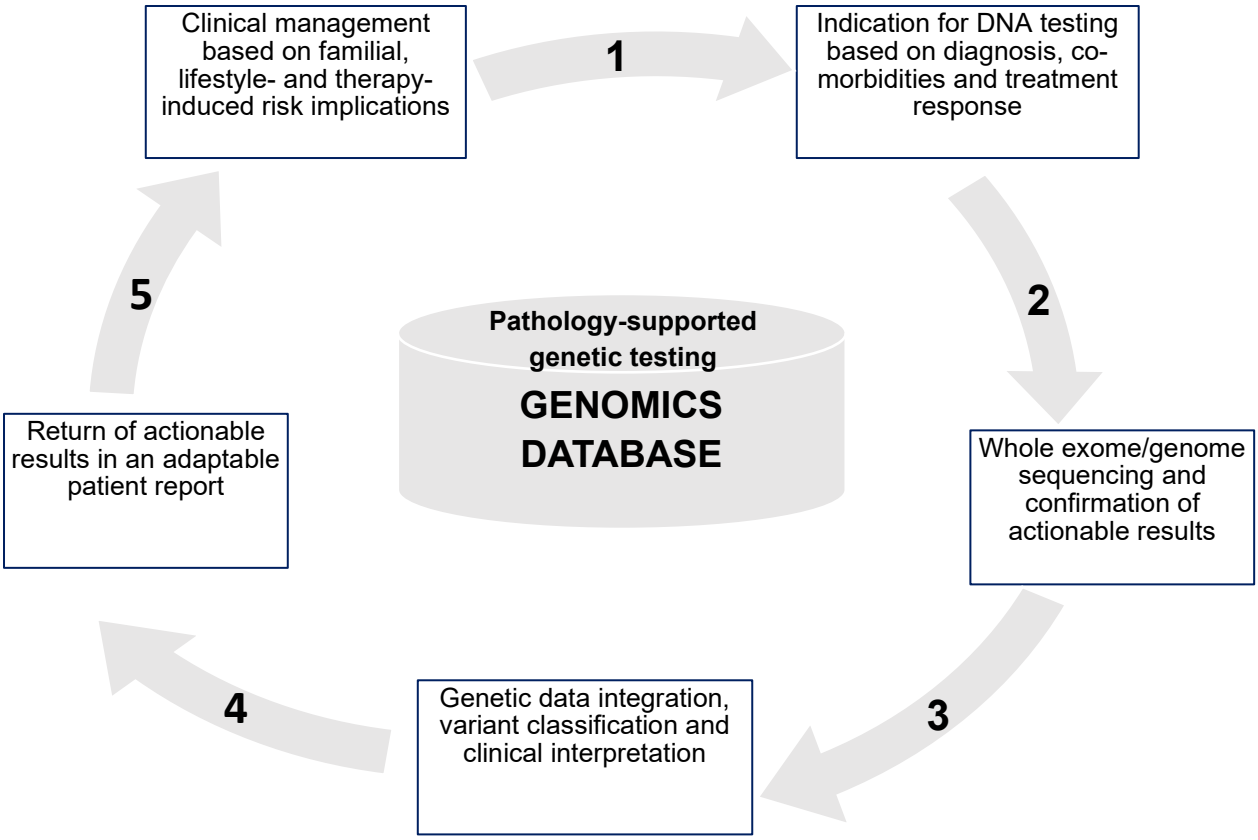


Figure 6.1: Process developed for integration of clinical, pathological and genetic results using the PSGT research translation platform. Treatment failure reported at any of the 5 steps are entered into the genomics database for re-interpretation of results in eligible cases.

Comparative WES and nanopore long-range sequencing of the *CDH1* gene using DNA samples of 7 patients with ILC confirmed the accuracy of the pocketsize MinION device used for the first time in our laboratory. The clinically enriched research database used in this study provides a secure online resource for research translation into adaptable reports, developed in an ongoing manner (Baatjes, 2018).

Future view

An alarming rate of vitamin D deficiency up to 40% of cases has been reported in breast cancer patients who developed metastasis (Maier et al., 2015). Given the high risk of metastasis to the bone that may become unmanageable, more effort should be made to target the bone microenvironment that plays an important role in cancer growth. Low vitamin D levels

remain largely undetected and untreated in high-risk breast cancer patients (Maier et al., 2015). Long-term follow-up of our study cohort is therefore warranted given the rich source of clinical information that have been captured in the genomics database generated on the PSGT platform. Development of bone metastasis is associated with interaction between bone and cancer cells that may be caused by endocrine changes contributing to release of bone-derived growth factors and cytokines promoting proliferation of cancer cells (Coleman, 2000).

This study highlighted the need to educate the public on the importance of maintaining normal body weight by engaging in regular physical activity. This will be helpful in release of trapped vitamin D in adipose tissue into circulation to treat vitamin D deficiency/insufficiency and reduce BMI in South African breast cancer patients. Nutrition (Weaver, 2013) and exercise support (Thomas et al., 2017) do not currently form an integral part of breast cancer risk management. Routine assessment of vitamin D levels is essential prior to aromatase inhibitor treatment of postmenopausal breast cancer patients to optimize bone health and quality of life.

In conclusion, the new insights gained as a result of this study relating to (i) lifestyle-triggered vitamin D deficiency (obesity, lack of outdoor physical activities), (ii) a genetic predisposition for osteoporosis (*VDR*), and (iii) tumour histopathology underlying inherited breast cancer and familial risk (*BRCA1/2*, *CDH1*), may be incorporated into the PSGT algorithm. This pharmaco-diagnostic platform facilitated clinical interpretation of WES in high-risk postmenopausal hormone receptor-positive breast cancer patients at increased risk of aromatase inhibitor-induced bone loss, recently found to be linked to the *CYP19A1* rs10046 AA genotype in the same study cohort. These findings support the use of WES alongside clinical assessments to assist clinicians in making the best treatment decision for their patients. WES may assist in identifying clinically relevant genetic risk factors for breast cancer development as well as treatment response with the application of a single test.

CHAPTER 7

REFERENCES

Aapro, M. S., & Coleman, R. E. (2012). Bone health management in patients with breast cancer: Current standards and emerging strategies. *The Breast*, *21*, 8–19. https://doi.org/10.1016/j.breast.2011.08.138

Abbas, S., Linseisen, J., Slanger, T., Kropp, S., Mutschelknauss, E. J., Flesch-Janys, D., & Chang-Claude, J. (2008). The Gc2 allele of the vitamin D binding protein is associated with a decreased postmenopausal breast cancer risk, independent of the vitamin D status. *Cancer Epidemiology Biomarkers & Prevention*, *17*(6), 1339–1343. https://doi.org/10.1158/1055-9965.EPI-08-0162

Abd-Elkader Abd-Elsalam, E., Ismaeil, N. A., Sibai, H., & Alsalam, A.-. (2015). Vitamin D receptor gene polymorphisms and breast cancer risk among postmenopausal Egyptian women. *Tumor Biology*, *36*(8), 6425–6431. https://doi.org/10.1007/s13277-015-3332-3

Abulkhair, O., Saadeddin, A., Makram, O., Gasmelseed, A., Pasha, T., Shehata, H., & Fakhoury, H. M. (2016). Vitamin D levels and breast cancer characteristics: Findings in patients from Saudi Arabia. *The Journal of Steroid Biochemistry and Molecular Biology*, *25*, 8–11. https://doi.org/10.1016/j.jsbmb.2015.11.003

Akinyemiju, T., Wiener, H., & Pisu, M. (2017). Cancer-related risk factors and incidence of major cancers by race, gender and region; analysis of the NIH-AARP diet and health study. *BMC Cancer*, *17*(1), 597. https://doi.org/10.1186/s12885-017-3557-1

Al-Daghri, N. M., Mohammed, A. K., Al-Attas, O. S., Ansari, M. G. A., Wani, K., Hussain, S. D., … Alokail, M. S. (2017). Vitamin D receptor gene polymorphisms modify cardiometabolic response to vitamin D supplementation in T2DM patients. *Scientific Reports*, *7*(1), 8280. https://doi.org/10.1038/s41598-017-08621-7

Alkhenizan, A., Mahmoud, A., Hussain, A., Gabr, A., Alsoghayer, S., & Eldali, A. (2017). The relationship between 25 (OH) D levels (vitamin D) and bone mineral density (BMD) in a Saudi population in a community-based Setting. *PLOS ONE*, *12*(1), e0169122. https://doi.org/10.1371/journal.pone.0169122

Allison, K. H. (2012). Molecular pathology of breast cancer: What a pathologist needs to know. *American Journal of Clinical Pathology*, *138*(6), 770–780. https://doi.org/10.1309/AJCPIV9IQ1MRQMOO

Allred, D. C. (2010). Ductal carcinoma in situ: Terminology, classification, and natural history. *Journal of the National Cancer Institute - Monographs*, (41), 134–138. https://doi.org/10.1093/jncimonographs/lgq035

Almesri, N., Das, N. S., Ali, M. E., Gumaa, K., & Giha, H. A. (2016). Independent associations of polymorphisms in vitamin D binding protein (GC) and vitamin D receptor (VDR) genes with obesity and plasma 25OHD3 levels demonstrate sex dimorphism. *Applied Physiology, Nutrition, and Metabolism*, *41*(4), 345–353. https://doi.org/10.1139/apnm-2015-0284

Amadori, D., Serra, P., Masalu, N., Pangan, A., Scarpi, E., Maria Bugingo, A., … Mercatali, L. (2017). Vitamin D receptor polymorphisms or serum levels as key drivers of breast cancer development? The question of the vitamin D pathway. *Oncotarget*, *8*(8), 13142–13156. https://doi.org/10.18632/oncotarget.14482

Anandabaskar, N., Selvarajan, S., & Kamalanathan, S. (2018). Vitamin D in health and disease – An update. *Journal of Young Pharmacists*, *10*(4), 381–387. https://doi.org/10.5530/jyp.2018.10.85

Anderson, L. N., Cotterchio, M., Cole, D. E. C., & Knight, J. A. (2011). Vitamin D-related genetic variants, interactions with vitamin D exposure, and breast cancer risk among caucasian women in Ontario. *Cancer Epidemiology Biomarkers and Prevention*, *20*(8),

1708–1717. https://doi.org/10.1158/1055-9965.EPI-11-0300

Andraos, C., Koorsen, G., Knight, J. C., & Bornman, L. (2011). Vitamin D receptor gene methylation is associated with ethnicity, tuberculosis, and TaqI polymorphism. *Human Immunology*, *72*(3), 262–268. https://doi.org/10.1016/j.humimm.2010.12.010

Arai, H., Miyamoto, K.-I., Taketani, Y., Yamamoto, H., Iemori, Y., Morita, K., … Takeda, E. (1997). A vitamin D receptor gene polymorphism in the translation initiation codon: Effect on protein activity and relation to bone mineral density in Japanese women. *Journal of Bone and Mineral Research*, *12*(6), 915–921. https://doi.org/10.1359/jbmr.1997.12.6.915

Ardawi, M.-S., Qari, M. H., Rouzi, A. A., Maimani, A. A., & Raddadi, R. M. (2010). Vitamin-D inadequacy in Saudi postmenopausal women with osteoporosis. *Bone*, *47*, S192. https://doi.org/10.1016/J.BONE.2010.04.455

Ashktorab, H., Azimi, H., Nickerson, M. L., Bass, S., Varma, S., & Brim, H. (2016). Targeted exome sequencing outcome variations of colorectal tumors within and across two sequencing platforms. *Next Generation, Sequencing & Applications*, *3*(1). https://doi.org/10.4172/2469-9853.1000123

Atoum, M., & Alzoughool, F. (2017). Vitamin D and breast cancer: latest evidence and future steps. *Breast Cancer: Basic and Clinical Research*, *11*. https://doi.org/10.1177/1178223417749816

Avenell, A., Bolland, M. J., & Grey, A. (2019). 25-Hydroxyvitamin D-Should labs be measuring it? *Annals of Clinical Biochemistry*, *56*(2), 188–189. https://doi.org/10.1177/0004563218796858

Baatjes, Karin J., Kotze, M. J., McCaul, M., & Conradie, M. (2019). Baseline bone health status in multi-ethnic South African postmenopausal breast cancer patients at initiation of aromatase inhibitor therapy: A descriptive study. *PLoS ONE*, *14*(4), e0214153.

https://doi.org/10.1371/journal.pone.0214153

Baatjes, Karin Jeanné. (2018). Bone health in post-menopausal patients with breast cancer treated with aromatase inhibitors: factors predicting the risk for osteoporosis . by. *Https://Scholar.Sun.Ac.Za/Handle/10019.1/105099*, 1–155.

Baatjes, Karin Jeanne, Conradie, M., Apffelstaedt, J. P., & Kotze, M. (2017). Pharmacogenetics of aromatase inhibitors in endocrine responsive breast cancer: lessons learnt from tamoxifen and CYP2D6 genotyping. *Anti-Cancer Agents in Medicinal Chemistry*, 1–9. https://doi.org/10.2174/1871521409666170412124226

Bai, X., Zhang, E., Ye, H., Nandakumar, V., Wang, Z., Chen, L., … Gao, J. (2014). PIK3CA and TP53 gene mutations in human breast cancer tumors frequently detected by ion torrent DNA sequencing. *PLoS ONE*, *9*(6), e99306. https://doi.org/10.1371/journal.pone.0099306

Balmaña, J., Díez, O., Rubio, I. T., & Cardoso, F. (2011). BRCA in breast cancer: ESMO clinical practice guidelines. *Annals of Oncology*, *22*(SUPPL. 6), 31–34. https://doi.org/10.1093/annonc/mdr373

Bandera Merchan, B., Morcillo, S., Martin-Nuñez, G., Tinahones, F. J., & Macías-González, M. (2017). The role of vitamin D and VDR in carcinogenesis: Through epidemiology and basic sciences. *Journal of Steroid Biochemistry and Molecular Biology*, *167*, 203–218. https://doi.org/10.1016/j.jsbmb.2016.11.020

Bane, A. L., Beck, J. C., Bleiweiss, I., Buys, S. S., Catalano, E., Daly, M. B., … O??Malley, F. P. (2007). BRCA2 Mutation-associated breast cancers exhibit a distinguishing phenotype based on morphology and molecular profiles from tissue microarrays. *The American Journal of Surgical Pathology*, *31*(1), 121–128. https://doi.org/10.1097/01.pas.0000213351.49767.0f

Barbarawi, M., Kheiri, B., Zayed, Y., Barbarawi, O., Dhillon, H., Swaid, B., … Manson, J. E.

(2019). Vitamin D supplementation and cardiovascular disease risks in more than 83 000 individuals in 21 randomized clinical trials. *JAMA Cardiology*, *4*(8), 765. https://doi.org/10.1001/jamacardio.2019.1870

Barroso-Sousa, R., & Metzger-Filho, O. (2016). Differences between invasive lobular and invasive ductal carcinoma of the breast: Results and therapeutic implications. *Therapeutic Advances in Medical Oncology*, *8*(4), 261–266. https://doi.org/10.1177/1758834016644156

Batai, K., Murphy, A. B., Shah, E., Ruden, M., Newsome, J., Agate, S., … Kittles, R. A. (2014). Common vitamin D pathway gene variants reveal contrasting effects on serum vitamin D levels in African Americans and European Americans. *Human Genetics*, *133*(11), 1395–1405. https://doi.org/10.1007/s00439-014-1472-y

Bauer, M., Bryce, J., & Hadji, P. (2012). Aromatase inhibitor-associated bone loss and its management with bisphosphonates in patients with breast cancer. *Breast Cancer: Targets and Therapy*, *4*, 91–101. https://doi.org/10.2147/BCTT.S29432

Behjati, S., & Tarpey, P. S. (2013). What is next generation sequencing? *Archives of Disease in Childhood - Education and Practice*, *98*(6), 236–238. https://doi.org/10.1136/archdischild-2013-304340

Benusiglio, P. R. (2017). CDH1 germline mutations: Different syndromes, same management? *Genetics in Medicine*, *19*(9), 965–966. https://doi.org/10.1038/gim.2017.24

Beral, V., & Million Women Study Collaborators. (2003). Breast cancer and hormone-replacement therapy in the million women study. *Lancet (London, England)*, *362*(9382), 419–427. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12927427

Berry, D., & Hyppönen, E. (2011). Determinants of vitamin D status: focus on genetic variations. *Current Opinion in Nephrology and Hypertension*, *20*(4), 331–336.

https://doi.org/10.1097/MNH.0b013e328346d6ba

Beysel, S., Eyerci, N., Pinarli, F. A., Apaydin, M., Kizilgul, M., Caliskan, M., … Cakal, E. (2018). VDR gene FokI polymorphism as a poor prognostic factor for papillary thyroid cancer. *Tumor Biology*, *40*(11), 101042831881176. https://doi.org/10.1177/1010428318811766

Bikle, D. D. (2014). Vitamin D metabolism, mechanism of action, and clinical applications. *Chemistry & Biology*, *21*(3), 319–329. https://doi.org/10.1016/j.chembiol.2013.12.016.Vitamin

Black, L. J., Lucas, R. M., Sherriff, J. L., Björn, L. O., & Bornman, J. F. (2017). In pursuit of vitamin D in plants. *Nutrients*, *9*(2), 1–9. https://doi.org/10.3390/nu9020136

Blackman, S., Dawnay, N., Ball, G., Stafford-Allen, B., Tribble, N., Rendell, P., … Wells, S. (2015). Developmental validation of the paraDNA® intelligence system - A novel approach to DNA profiling. *Forensic Science International: Genetics*, *17*, 137–148. https://doi.org/10.1016/j.fsigen.2015.04.009

Blanco-Gómez, A., Castillo-Lluva, S., del Mar Sáez-Freire, M., Hontecillas-Prieto, L., Mao, J. H., Castellanos-Martín, A., & Pérez-Losada, J. (2016). Missing heritability of complex diseases: Enlightenment by genetic variants from intermediate phenotypes. *BioEssays*, *38*(7), 664–673. https://doi.org/10.1002/bies.201600084

Bombard, Y., Brothers, K. B., Fitzgerald-Butt, S., Garrison, N. A., Jamal, L., James, C. A., … Levy, H. P. (2019). The responsibility to recontact research participants after reinterpretation of genetic and genomic research results. *American Journal of Human Genetics*, *104*(4), 578–595. https://doi.org/10.1016/j.ajhg.2019.02.025

Bouvard, B., Hoppé, E., Soulié, P., Georgin-Mege, M., Jadaud, E., Abadie-Lacourtoisie, S., … Legrand, E. (2012). High prevalence of vertebral fractures in women with breast cancer starting aromatase inhibitor therapy. *Annals of Oncology / ESMO*, *23*(5), 1151–

1156. https://doi.org/10.1093/annonc/mdr356

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *68*(6), 394–424. https://doi.org/10.3322/caac.21492

Brewer, H. R., Jones, M. E., Schoemaker, M. J., Ashworth, A., & Swerdlow, A. J. (2017). Family history and risk of breast cancer: an analysis accounting for family structure. *Breast Cancer Research and Treatment*, *165*(1), 193–200. https://doi.org/10.1007/s10549-017-4325-2

Brinton, L. A., Figueroa, J. D., Awuah, B., Yarney, J., Wiafe, S., Wood, S. N., … Clegg-Lamptey, J. N. (2014). Breast cancer in Sub-saharan Africa: Opportunities for prevention. *Breast Cancer Research and Treatment*, *144*(3), 467–478. https://doi.org/10.1007/s10549-014-2868-z

Brouwer-Brolsma, E. M., Bischoff-Ferrari, H. A., Bouillon, R., Feskens, E. J. M., Gallagher, C. J., Hypponen, E., … de Groot, L. C. P. G. M. (2013). Vitamin D: do we get enough? *Osteoporosis International*, *24*(5), 1567–1577. https://doi.org/10.1007/s00198-012-2231-3

Cabanillas, R., Diñeiro, M., Castillo, D., Pruneda, P. C., Penas, C., Cifuentes, G. A., … Cadiñanos, J. (2017). A novel molecular diagnostics platform for somatic and germline precision oncology. *Molecular Genetics & Genomic Medicine*, *5*(4), 336–359. https://doi.org/10.1002/mgg3.291

Caldeira, J. R. F., Prando, É. C., Quevedo, F. C., Neto, F. A. M., Rainho, C. A., & Rogatto, S. R. (2006). CDH1promoter hypermethylation and E-cadherin protein expression in infiltrating breast cancer. *BMC Cancer*, *6*(1), 48. https://doi.org/10.1186/1471-2407-6-48

Carpenter, T. O., Zhang, J. H., Parra, E., Ellis, B. K., Simpson, C., Lee, W. M., … Cole, D. E.

C. (2013). Vitamin D binding protein is a key determinant of 25-hydroxyvitamin D levels in infants and toddlers. *Journal of Bone and Mineral Research*, *28*(1), 213–221. https://doi.org/10.1002/jbmr.1735

Chakravorty, S., & Hegde, M. (2017). Gene and variant annotation for mendelian disorders in the era of advanced sequencing technologies. *Annual Review of Genomics and Human Genetics*, *18*, 229–256. https://doi.org/10.1146/annurev-genom-083115

Chandler, P., Tobias, D., Wang, L., Smith-Warner, S., Chasman, D., Rose, L., … Sesso, H. (2018). Association between Vitamin D genetic risk score and cancer risk in a large cohort of u.s. women. *Nutrients*, *10*(1), 55. https://doi.org/10.3390/nu10010055

Channel, B. H. (2018). Breast cancer and oestrogen - Better Health Channel. Retrieved December 17, 2019, from https://www.betterhealth.vic.gov.au/health/conditionsandtreatments/breast-cancer-and-oestrogen

Chen, F., Chen, G. K., Stram, D. O., Millikan, R. C., Ambrosone, C. B., John, E. M., … Haiman, C. A. (2013). A genome-wide association study of breast cancer in women of African ancestry. *Human Genetics*, *132*(1), 39–48. https://doi.org/10.1007/s00439-012-1214-y

Chen, H., Wu, J., Zhang, Z., Tang, Y., Li, X., Liu, S., … Li, X. (2018). Association between *BRCA* status and triple-negative breast cancer: A meta-analysis. *Frontiers in Pharmacology*, *9*, 909. https://doi.org/10.3389/fphar.2018.00909

Christakos, S., Dhawan, P., Verstuyf, A., Verlinden, L., & Carmeliet, G. (2016). Vitamin D: Metabolism, molecular mechanism of action, and pleiotropic effects. *Physiological Reviews*, *96*(1), 365–408. https://doi.org/10.1152/physrev.00014.2015

Coleman, R. E. (2000). Management of bone metastases. *The Oncologist*, *5*(6), 463–470. https://doi.org/10.1634/theoncologist.5-6-463

Collins, F. S. (2004). What we do and don't know about "race", "ethnicity", genetics and health at the dawn of the genome era. *Nature Genetics*, *36*(S11), S13–S15. https://doi.org/10.1038/ng1436

Cooper, D. N., Krawczak, M., Polychronakos, C., Tyler-Smith, C., & Kehrer-Sawatzki, H. (2013). Where genotype is not predictive of phenotype: Towards an understanding of the molecular basis of reduced penetrance in human inherited disease. *Human Genetics* (Vol. 132). https://doi.org/10.1007/s00439-013-1331-2

Corso, G., Intra, M., Chiara Trentin, Veronesi, P., & Galimberti, V. (2016). CDH1 germline mutations and hereditary lobular breast cancer. *Familial Cancer*, *15*, 215–219. https://doi.org/10.1007/s10689-016-9869-5

Coussens, A. K., Naude, C. E., Goliath, R., Chaplin, G., Wilkinson, R. J., & Jablonski, N. G. (2015). High-dose vitamin D_3 reduces deficiency caused by low UVB exposure and limits HIV-1 replication in urban Southern Africans. *Proceedings of the National Academy of Sciences of the United States of America*, (19), 1–6. https://doi.org/10.1073/pnas.1500909112

Cristofanilli, M., & Fortina, P. (2013). Circulating tumor DNA to monitor metastatic breast cancer. *New England Journal of Medicine*. https://doi.org/10.1056/NEJMc1306040

Curtis, C., Shah, S. P., Chin, S.-F., Turashvili, G., Rueda, O. M., Dunning, M. J., … Aparicio, S. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*, *486*(7403), 346–352. https://doi.org/10.1038/nature10983

Damineni, S., Rao, V. R., Kumar, S., Ravuri, R. R., Kagitha, S., Dunna, N. R., … Satti, V. (2014). Germline mutations of TP53 gene in breast cancer. *Tumor Biology*, *35*(9), 9219–9227. https://doi.org/10.1007/s13277-014-2176-6

Datta, P., Philipsen, P. A., Olsen, P., Petersen, B., Andersen, J. D., Morling, N., & Wulf, H. C. (2019). Pigment genes not skin pigmentation affect UVB-induced vitamin D.

Photochemical and Photobiological Sciences, *18*(2), 448–458. https://doi.org/10.1039/C8PP00320C

Dawson-hughes, B., & Harris, S. S. (2002). Calcium intake influences the association of protein intake with rates of bone loss in elderly men and women. *The American Journal of Clinical Nutrition*, (75)4, 773–779. https://doi.org/10.3109/10408363.2014.997930

Dawson, S.J., Tsui, D. W., Murtaza, M., Biggs, H., Rueda, O. M., Chin, S.-F., … Rosenfeld, N. (2013). Analysis of circulating tumor DNA to monitor metastatic breast cancer. *New England Journal of Medicine*, *368*(13), 1199–1209. https://doi.org/10.1056/NEJMoa1213261

Dewey, F. E., Chen, R., Cordero, S. P., Ormond, K. E., Caleshu, C., Karczewski, K. J., … Ashley, E. A. (2011). Phased whole-genome genetic risk in a family quartet using a major allele reference sequence. *PLoS Genetics*, *7*(9), e1002280. https://doi.org/10.1371/journal.pgen.1002280

Dickens, C., Duarte, R., Zietsman, A., Cubasch, H., Kellett, P., Schüuz, J., … McCormack, V. (2014). Racial comparison of receptor-defined breast cancer in Southern African women: Subtype prevalence and age - Incidence analysis of nationwide cancer registry data. *Cancer Epidemiology Biomarkers and Prevention*, *23*(11), 2311–2321. https://doi.org/10.1158/1055-9965.EPI-14-0603

Dimitrakopoulou, V. I., Tsilidis, K. K., Haycock, P. C., Dimou, N. L., Al-Dabhani, K., Martin, R. M., … Schildkraut, J. M. (2017). Circulating vitamin D concentration and risk of seven cancers: Mendelian randomisation study. *BMJ (Clinical Research Ed.)*, *359*, j4761. https://doi.org/10.1136/bmj.j4761

Dine, J., & Deng, C.-X. (2013). Mouse models of *BRCA1* and their application to breast cancer research. *Cancer Metastasis Rev.*, *32*, 25–37. https://doi.org/10.1007/s10555-012-9403-7

Dixon, J. M. (2014). Endocrine resistance in breast cancer. *New Journal of Science*, *2014*(10), 1–27. https://doi.org/10.1155/2014/390618

Donnelly, M. P., Paschou, P., Grigorenko, E., Gurwitz, D., Barta, C., Lu, R.-B., … Kidd, K. K. (2012). A global view of the *OCA2-HERC2* region and pigmentation. *Human Genetics*, *131*(5), 683–696. https://doi.org/10.1007/s00439-011-1110-x

Edwards, B. J., Raisch, D. W., Shankaran, V., McKoy, J. M., Gradishar, W., Bunta, A. D., … Guise, T. A. (2011). Cancer therapy associated bone loss: implications for hip fractures in mid-life women with breast cancer. *Clinical Cancer Research*, *17*(3), 560–568. https://doi.org/10.1158/1078-0432.CCR-10-1595

El-Shorbagy, S. EL. (2017). Prognostic impact of 25-hydroxyvitamin D levels in Egyptian patients with breast cancer. *Journal of Cancer Science & Therapy*, *9*(6), 496–502. https://doi.org/10.4172/1948-5956.1000466

Elizondo-Montemayor, L., Castillo, E. C., Rodríguez-López, C., Villarreal-Calderón, J. R., Gómez-Carmona, M., Tenorio-Martínez, S., … García-Rivas, G. (2017). Seasonal variation in vitamin D in association with age, inflammatory cytokines, anthropometric parameters, and lifestyle factors in older adults. *Mediators of Inflammation*, *2017*, 1–14. https://doi.org/10.1155/2017/5719461

Engelman, C. D., Fingerlin, T. E., Langefeld, C. D., Hicks, P. J., Rich, S. S., Wagenknecht, L. E., … Norris, J. M. (2008). Genetic and environmental determinants of 25-hydroxyvitamin D and 1,25-dihydroxyvitamin D levels in Hispanic and African Americans. *Journal of Clinical Endocrinology and Metabolism*, *93*(9), 3381–3388. https://doi.org/10.1210/jc.2007-2702

Farmer, P., Frenk, J., Knaul, F. M., Shulman, L. N., Alleyne, G., Armstrong, L., … Seffrin, J. R. (2010). Expansion of cancer care and control in countries of low and middle income: A call to action. *The Lancet*, *376*(9747), 1186–1193. https://doi.org/10.1016/S0140-

6736(10)61152-X

Feng, Y., Spezia, M., Huang, S., Yuan, C., Zeng, Z., Zhang, L., … Ren, G. (2018). Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes and Diseases*, *5*(2), 77–106. https://doi.org/10.1016/j.gendis.2018.05.001

Ferlay, J, Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D. M., Piñeros, M., … Bray, F. (2018). Global cancer observatory: Cancer today. Lyon, France: International Agency for Research on Cancer. *Https://Gco.Iarc.Fr/Today*, *144*(8), 1941–1953. https://doi.org/10.1002/ijc.31937

Ferlay, Jacques, Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., … Bray, F. (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, *136*(5), E359–E386. https://doi.org/10.1002/ijc.29210

Fetahu, I. S., Höbaus, J., & Kállay, E. (2014). Vitamin D and the epigenome. *Frontiers in Physiology*, *5*, 1–12. https://doi.org/10.3389/fphys.2014.00164

Figueiredo, J., Melo, S., Carneiro, P., Moreira, A. M., Fernandes, M. S., Ribeiro, A. S., … Seruca, R. (2019). Clinical spectrum and pleiotropic nature of *CDH1* germline mutations. *Journal of Medical Genetics*, *56*(4), 199–208. https://doi.org/10.1136/jmedgenet-2018-105807

Fourie, J., Motswaledi, M. H., Khammissa, R. A. G., Lemmer, J., Ballyram, R., & Feller, L. (2018). The biological activities of vitamin D and its receptor in relation to calcium and bone homeostasis, cancer, immune and cardiovascular systems, skin biology, and oral health. *BioMed Research International*, *2018*, 1–9. https://doi.org/10.1155/2018/9276380

Fraser, W. D., & Milan, A. M. (2013). Vitamin D assays: Past and present debates,

difficulties, and developments. *Calcified Tissue International*, *92*(2), 118–127. https://doi.org/10.1007/s00223-012-9693-3

Frey, J. D., Salibian, A. A., Schnabel, F. R., Choi, M., & Karp, N. S. (2017). Non-*BRCA1/2* breast cancer susceptibility genes: A new frontier with clinical consequences for plastic surgeons. *Plastic and Reconstructive Surgery. Global Open*, *5*(11), e1564. https://doi.org/10.1097/GOX.0000000000001564

Garland, C. F., Garland, F. C., Gorham, E. D., Lipkin, M., Newmark, H., Mohr, S. B., & Holick, M. F. (2006). The role of vitamin D in cancer prevention. *American Journal of Public Health*, *96*(2), 252–261. https://doi.org/10.2105/AJPH.2004.045260

Garland, C. F., Gorham, E. D., Mohr, B., & Garland, F. C. (2009). Vitamin D for cancer prevention: Global perspective. *Annals of Epidemiology*, *19*(7), 468–483. https://doi.org/10.1016/j.annepidem.2009.03.021

Georgios Papadakis, Ioannis Keramidas, K. K., Theodora Pappa, Vassiliki Villiotou, Eleni Triantafillou, A. D., & Athanasia Tertipi, V. K. and A. P. (2015). Seasonal variation of serum vitamin D among greek female patients with osteoporosis. *In Vivo*, *29*, 409–414. Retrieved from http://iv.iiarjournals.org/content/29/3/409.full.pdf

Golemis, E. A., Scheet, P., Beck, T. N., Scolnick, E. M., Hunter, D. J., Hawk, E., & Hopkins, N. (2018). Molecular mechanisms of the preventable causes of cancer in the United States. *Genes and Development*, *32*(13–14), 868–902. https://doi.org/10.1101/gad.314849.118

Goodwin, S., McPherson, J. D., & McCombie, W. R. (2016). Coming of age: Ten years of next-generation sequencing technologies. *Nature Reviews Genetics*, *17*(6), 333–351. https://doi.org/10.1038/nrg.2016.49

Gracia-Aznarez, F. J., Fernandez, V., Pita, G., Peterlongo, P., Dominguez, O., de la Hoya, M., … Benitez, J. (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*(2). https://doi.org/10.1371/journal.pone.0055681

Grant, W. B., & Boucher, B. J. (2017). Randomized controlled trials of vitamin D and cancer incidence: A modeling study. *PLoS ONE* 12(5), e0176448. doi:10.1371/journal.pone.0176448

Graziano, S., Johnston, R., Deng, O., Zhang, J., & Gonzalo, S. (2016). Vitamin D/vitamin D receptor axis regulates DNA repair during oncogene-induced senescence. *Oncogene*, *2*(74), 5362–5376. https://doi.org/10.1126/scisignal.274pe36.Insulin

Greif, J. M., Pezzi, C. M., Klimberg, S. V., Bailey, L., & Zuraek, M. (2012). Gender differences in breast cancer: Analysis of 13,000 breast cancers in men from the national cancer data base. *Annals of Surgical Oncology*, *19*(10), 3199–3204. https://doi.org/10.1245/s10434-012-2479-z

Grotsky, D. A., Gonzalez-Suarez, I., Novell, A., Neumann, M. A., Yaddanapudi, S. C., Croke, M., … Gonzalo, S. (2013). *BRCA1* loss activates cathepsin L-mediated degradation of *53BP1* in breast cancer cells. *The Journal of Cell Biology*, *200*(2), 187–202. https://doi.org/10.1083/jcb.201204053

Haarburger, D., Hoffman, M., Erasmus, R. T., & Pillay, T. S. (2009). Relationship between Vitamin D, calcium and PTH in Cape Town. *Journal of Clinical Pathology*, 62(6), 567-9 https://doi.org/10.1136/jcp.2008.062877

Hadgu, E., Seifu, D., Tigneh, W., Bokretsion, Y., Bekele, A., Abebe, M., … Karlsson, M. G. (2018). Breast cancer in Ethiopia: Evidence for geographic difference in the distribution of molecular subtypes in Africa. *BMC Women's Health*, *18*(1), 1–8. https://doi.org/10.1186/s12905-018-0531-2

Han, M. R., Zheng, W., Cai, Q., Gao, Y. T., Zheng, Y., Bolla, M. K., … Long, J. (2017). Evaluating genetic variants associated with breast cancer risk in high and moderate-

penetrance genes in Asians. *Carcinogenesis*, *38*(5), 511–518. https://doi.org/10.1093/carcin/bgx010

Hanby AM and Hughes TA. (2008). In situ and invasive lobular neoplasia of the breast. *Histopathology*, *52*, 58–66. https://doi.org/10.1111/j.1365-2559.2007.02891.x

Haque, R., Ahmed, S. A., Fisher, A., Avila, C. C., Shi, J., Guo, A., … Schottinger, J. E. (2012). Effectiveness of aromatase inhibitors and tamoxifen in reducing subsequent breast cancer. *Cancer Medicine*, *1*(3), 318–327. https://doi.org/10.1002/cam4.37

Hatse, S., Lambrechts, D., Verstuyf, A., Smeets, A., Brouwers, B., Vandorpe, T., … Wildiers, H. (2012). Vitamin D status at breast cancer diagnosis: Correlation with tumor characteristics, disease outcome, and genetic determinants of vitamin D insufficiency. *Carcinogenesis*, *33*(7), 1319–1326. https://doi.org/10.1093/carcin/bgs187

Haussler, M. R., Whitfield, G. K., Kaneko, I., Haussler, C. A., Hsieh, D., Hsieh, J. C., & Jurutka, P. W. (2013). Molecular mechanisms of vitamin D action. *Calcified Tissue International*, *92*(2), 77–98. https://doi.org/10.1007/s00223-012-9619-0

He, W., Liu, M., Huang, X., Qing, Z., & Gao, W. (2015). The influence of vitamin D receptor genetic variants on bone mineral density and osteoporosis in Chinese postmenopausal women. *Disease Markers*, *2015*, 1–7. https://doi.org/10.1155/2015/760313

Health, N. I. of. (2013). Sex hormones and breast cancer risk in premenopausal women: collaborative reanalysis of seven prospective studies. *Lancet Oncology*, *14*(10), 1009–1019. https://doi.org/10.1016/S1470-2045(13)70301-2.Sex

Hengist, A., Perkin, O., Gonzalez, J. T., Betts, J. A., Hewison, M., & Manolopoulos, K. N. (2019). Mobilising vitamin D from adipose tissue : The potential impact of exercise. *Nutrition Bulletin*, *3*, 25–35. https://doi.org/10.1111/nbu.12369

Hodges, E., Xuan, Z., Balija, V., Kramer, M., Molla, M. N., Smith, S. W., … McCombie, W. R.

(2007). Genome-wide in situ exon capture for selective resequencing. *Nature Genetics*, *39*(12), 1522–1527. https://doi.org/10.1038/ng.2007.42

Holick, M. F. (2009). Vitamin D status: Measurement, interpretation and clinical application. *Annals of Epidemiology*, *19*(2), 73–78. https://doi.org/10.1016/j.annepidem.2007.12.001.VITAMIN

Hossein-nezhad, A., Spira, A., & Holick, M. F. (2013). Influence of vitamin D Status and vitamin D_3 supplementation on genome wide expression of white blood cells: A randomized double-blind clinical trial. *PLoS ONE*, *8*(3). https://doi.org/10.1371/journal.pone.0058725

Huang, X., Shu, C., Chen, L., & Yao, B. (2018). Impact of sex, body mass index and initial pathologic diagnosis age on the incidence and prognosis of different types of cancer. *Oncology Reports*, *40*(3), 1359–1369. https://doi.org/10.3892/or.2018.6529

Ingle, J. N., Buzdar, A. U., Schaid, D. J., Goetz, M. P., Batzler, A., Robson, M. E., … Weinshilboum, R. M. (2010). Variation in anastrozole metabolism and pharmacodynamics in women with early breast cancer. *Cancer Research*, *70*(8), 3278–3286. https://doi.org/10.1158/0008-5472.CAN-09-3024

Ismail, A., El-Awady, R., Mohamed, G., Hussein, M., & Ramadan, S. S. (2018). Prognostic significance of serum vitamin D levels in Egyptian females with breast cancer. *Asian Pacific Journal of Cancer Prevention*, *19*(2), 571–576. https://doi.org/10.22034/APJCP.2018.19.2.571

Jeon, S.-M., & Shin, E.-A. (2018). Exploring vitamin D metabolism and function in cancer. *Experimental & Molecular Medicine*, *50*(4), 20. https://doi.org/10.1038/s12276-018-0038-9

Jiang, T., Li, L., Wang, Y., Zhao, C., Yang, J., Ma, D., … Yang, J. (2016). The association between genetic polymorphism *rs703842* in *CYP27B1* and multiple sclerosis: A meta-

analysis. *Medicine*, *95*(19), e3612. https://doi.org/10.1097/MD.0000000000003612

Jiang, X., Dimou, N. L., Al-Dabhani, K., Lewis, S. J., Martin, R. M., Haycock, P. C., … Tsilidis, K. K. (2018). Circulating vitamin D concentrations and risk of breast and prostate cancer: a Mendelian randomization study. *International Journal of Epidemiology*. https://doi.org/10.1093/ije/dyy284

Joffe, M., Ayeni, O., Norris, S. A., McCormack, V. A., Ruff, P., Das, I., … Cubasch, H. (2018). Barriers to early presentation of breast cancer among women in Soweto, South Africa. *PLoS ONE*, *13*(2), 1–16. https://doi.org/10.1371/journal.pone.0192071

Jones, G., Prosser, D. E., & Kaufmann, M. (2012). 25-Hydroxyvitamin D-24-hydroxylase (CYP24A1): Its important role in the degradation of vitamin D. *Archives of Biochemistry and Biophysics*, *523*(1), 9–18. https://doi.org/10.1016/J.ABB.2011.11.003

Jones, P. A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, *13*. https://doi.org/10.1038/nrg3230

Jorde, R., Schirmer, H., Wilsgaard, T., Mathiesen, E. B., Njølstad, I., Løchen, M. L., … Grimnes, G. (2015). The DBP phenotype *Gc-1f/Gc-1f* is associated with reduced risk of cancer. The Tromsø study. *PLoS ONE*, *10*(5), 1–11. https://doi.org/10.1371/journal.pone.0126359

Jorde, R., Sneve, M., Hutchinson, M., Emaus, N., Figenschau, Y., & Grimnes, G. (2010). Tracking of serum 25-hydroxyvitamin D levels during 14 years in a population-based study and during 12 months in an intervention study. *American Journal of Epidemiology*, *171*(8), 903–908. https://doi.org/10.1093/aje/kwq005

Jung, S., Wang, M., Anderson, K., Baglietto, L., Bergkvist, L., Bernstein, L., … Smith-Warner, S. A. (2016). Alcohol consumption and breast cancer risk by estrogen receptor status: In a pooled analysis of 20 studies. *International Journal of Epidemiology*, *45*(3), 916–928. https://doi.org/10.1093/ije/dyv156

Kamińska, M., Ciszewski, T., Łopacka-Szatan, K., Miotła, P., & Starosławska, E. (2015). Breast cancer risk factors. *Przeglad Menopauzalny*, *14*(3), 196–202. https://doi.org/10.5114/pm.2015.54346

Karen L. Maughan, Mark A. Lutterbie, and P. S. H. (2010). Treatment of breast cancer. *American Family Physician*, *81*(11), 1339-46. https://doi.org/10.1016/S0140-6736(73)91254-3

Karlic, H., & Varga, F. (2011). Impact of vitamin D metabolism on clinical epigenetics. *Clin Epigenet*, *2*, 55–61. https://doi.org/10.1007/s13148-011-0021-y

Kazemian, E., Akbari, M. E., Moradi, N., Gharibzadeh, S., Mondul, A. M., Jamshidi-Naeini, Y., … Rozek, L. S. (2019). Vitamin D receptor genetic variation and cancer biomarkers among breast cancer patients supplemented with vitamin D_3: A single-arm non-randomized before and after trial. *Nutrients*, *11*(6). https://doi.org/10.3390/nu11061264

Kazemian, E., Amouzegar, A., Akbari, M. E., Moradi, N., Gharibzadeh, S., Jamshidi-Naeini, Y., … Davoodi, S. H. (2019). Vitamin D receptor gene polymorphisms affecting changes in visceral fat, waist circumference and lipid profile in breast cancer survivors supplemented with vitamin D_3. *Lipids in Health and Disease*, *18*(1), 161. https://doi.org/10.1186/s12944-019-1100-x

Kennel, K. a, Drake, M. T., & Hurley, D. L. (2010). Vitamin D deficiency in adults: when to test and how to treat. *Mayo Clinic Proceedings*, *85*(8), 752–758. https://doi.org/10.4065/mcp.2010.0138

Key, T., Appleby, P., Barnes, I., Reeves, G., & Endogenous Hormones and Breast Cancer Collaborative Group. (2002). Endogenous sex hormones and breast cancer in postmenopausal women: Reanalysis of nine prospective studies. *Journal of the National Cancer Institute*, *94*(8), 606–616. http://www.ncbi.nlm.nih.gov/pubmed/11959894

Khammissa, R. A. G., Fourie, J., Motswaledi, M. H., Ballyram, R., Lemmer, J., & Feller, L. (2018). The biological activities of vitamin D and its receptor in relation to calcium and bone homeostasis, cancer, immune and cardiovascular systems, skin biology, and oral health. *BioMed Research International*, *2018*, 1–9. https://doi.org/10.1155/2018/9276380

Khan, Q. J., Reddy, P. S., Kimler, B. F., Sharma, P., Baxa, S. E., O'Dea, A. P., … Fabian, C. J. (2010). Effect of vitamin D supplementation on serum 25-hydroxy vitamin D levels, joint pain, and fatigue in women starting adjuvant letrozole treatment for breast cancer. *Breast Cancer Research and Treatment*, *119*(1), 111–118. https://doi.org/10.1007/s10549-009-0495-x

Khodarahmi, M., & Azadbakht, L. (2014). The association between different kinds of fat intake and breast cancer risk in women. *International Journal of Preventive Medicine*, *5*(1), 6–15. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24554986

Kim, H. J., Lee, Y. M., Ko, B. S., Lee, J. W., Yu, J. H., Son, B. H., … Ahn, S. H. (2011). Vitamin D deficiency is correlated with poor outcomes in patients with luminal-type breast cancer. *Annals of Surgical Oncology*, *18*(7), 1830–1836. https://doi.org/10.1245/s10434-010-1465-6

Kim, S. W., Lee, J. M., Ha, J. H., Kang, H. H., Rhee, C. K., Kim, J. W., … Lee, S. H. (2015). Association between vitamin D receptor polymorphisms and osteoporosis in patients with COPD. *International Journal of Chronic Obstructive Pulmonary Disease*, *10*, 1809–1817. https://doi.org/10.2147/COPD.S91576

Kimble, R. B., Srivastava, S., Patrick Ross, F., Matayoshi, A., & Pacifici, R. (1996). Estrogen deficiency increases the ability of stromal cells to support murine osteoclastogenesis via an interleukin-1-and tumor necrosis factor- mediated stimulation of macrophage colony-stimulating factor production. *Journal of Biological Chemistry*, *271*(46), 28890–28897. https://doi.org/10.1074/jbc.271.46.28890

Kotepui, M. (2016). Diet and risk of breast cancer. *Contemporary Oncology*, *20*(1), 13–19. https://doi.org/10.5114/wo.2014.40560

Kotze MJ, van der Merwe, Mumpunye L, Grant KA, Peeters AV, French DJ. A rapid point-of-care test for detection of pathogenic *BRCA1/2* founder mutations: pharmacogenetic evaluation of South African breast cancer patients selected by tumour molecular subtype. *Genetics Research* (in press).

Kotze, M. (2016). Application of advanced molecular technology in the diagnosis and management of genetic disorders in South Africa. *South African Medical Journal, 106*(6), s114–s118. https://doi.org/10.7196/SAMJ.2016.v106i6.11012

Kotze, M. J., Lückhoff, H. K., Peeters, A. V, Baatjes, K., Schoeman, M., van der Merwe, L., … Schneider, J. W. (2015). Genomic medicine and risk prediction across the disease spectrum. *Critical Reviews in Clinical Laboratory Sciences*, *52*(3), 120–137. https://doi.org/10.3109/10408363.2014.997930

Kotze, M. J., van Velden, D. P., Botha, K., Badenhorst, C. H., Avenant, H., van Rensburg, S. J., & Cronjé, F. J. (2013). Pathology-supported genetic testing directed at shared disease pathways for optimized health in later life. *Personalized Medicine*, *10*(5), 497–507. https://doi.org/10.2217/pme.13.43

Krishnan, A. V., & Feldman, D. (2011). Mechanisms of the anti-cancer and anti-inflammatory actions of vitamin D. *Annual Review of Pharmacology and Toxicology*, *51*(1), 311–336. https://doi.org/10.1146/annurev-pharmtox-010510-100611

Krishnan, A. V., Swami, S., Peng, L., Wang, J., Moreno, J., & Feldman, D. (2010). Tissue-selective regulation of aromatase expression by calcitriol: Implications for breast cancer therapy. *Endocrinology*, *151*(1), 32–42. https://doi.org/10.1210/en.2009-0855

Lakhani, S. R., Khanna, K. K., & Chenevix-Trench, G. (2010). Are estrogen receptor-positive breast cancers in *BRCA1* mutation carriers sporadic? *Breast Cancer Research*, *12*(2),

7–8. https://doi.org/10.1186/bcr2483

Lakhani SR, Ellis IO, Schnitt SJ, Tan PH, van de V. M. (2012). *WHO classification of tumours of the breast*. (4th edition). Lyon: International Agency for Research on Cancer (IARC). Retrieved from https://shop.iarc.fr/products/classification-of-tumours-of-the-breast?variant=7964963504176

Lappe, J. M., & Heaney, R. P. (2012). Why randomized controlled trials of calcium and vitamin D sometimes fail. *Dermato-Endocrinology*, *4*(2), 95–100. https://doi.org/10.4161/derm.19833

Lee, S.-W., Chuang, T.-Y., Huang, H.-H., Liu, C.-W., Kao, Y.-H., & Shih-Hsin Wu, L. (2016). VDR and VDBP genes polymorphisms associated with susceptibility to tuberculosis in a Han Taiwanese population. *Immunology and Infection*, *49*, 783–787. https://doi.org/10.1016/j.jmii.2015.12.008

Leggett, R. M., & Clark, M. D. (2017). A world of opportunities with nanopore sequencing. *Journal of Experimental Botany*, *68*(20), 5419–5429. https://doi.org/10.1093/jxb/erx289

Li, C I, Uribe, D. J. and, & Daling, J. R. (2005). Clinical characteristics of different histologic types of breast cancer. *British Journal of Cancer*, *93*, 1046–1052. https://doi.org/10.1038/sj.bjc.6602787

Li, Christopher I., Anderson, B. O., Daling, J. R., & Moe, R. E. (2003). Trends in incidence rates of invasive lobular and ductal breast carcinoma. *Journal of the American Medical Association*, *289*(11), 1421. https://doi.org/10.1001/jama.289.11.1421

Li, W., Richter, R. A., Jung, Y., Zhu, Q., & Li, R. W. (2016). Web-based bioinformatics workflows for end-to-end RNA-seq data computation and analysis in agricultural animal species. *BMC Genomics*, *17*(1), 761. https://doi.org/10.1186/s12864-016-3118-z

Liu, J., Sun, X., Qin, S., Wang, H., DU, N., Li, Y., … Ren, H. (2016). *CDH1* promoter

methylation correlates with decreased gene expression and poor prognosis in patients with breast cancer. *Oncology Letters*, *11*(4), 2635–2643. https://doi.org/10.3892/ol.2016.4274

Liu, J. Y., Chen, T. J., & Hwang, S. J. (2016). The risk of breast cancer in women using menopausal hormone replacement therapy in Taiwan. *International Journal of Environmental Research and Public Health*, *13*(5). https://doi.org/10.3390/ijerph13050482

Lombard, Z., Dalton, D. L., Venter, P. A., Williams, R. C., & Bornman, L. (2006). Association of *HLA-DR, -DQ*, and vitamin D receptor alleles and haplotypes with tuberculosis in the Venda of South Africa. *Human Immunology*, *67*(8), 643–654. https://doi.org/10.1016/j.humimm.2006.04.008

Lopes, N., Carvalho, J., Durães, C., Sousa, B., Gomes, M., Costa, J. L., … Schmitt, F. (2012). 1Alpha,25-dihydroxyvitamin D_3 induces de novo E-cadherin expression in triple-negative breast cancer cells by CDH1-promoter demethylation. *Anticancer Research*, *32*(1), 249–257. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22213313

Lopes, N., Sousa, B., Martins, D., Gomes, M., Vieira, D., Veronese, L. A., … Schmitt, F. (2010). Alterations in vitamin D signalling and metabolic pathways in breast cancer progression: A study of *VDR, CYP27B1* and *CYP24A1* expression in benign and malignant breast lesions vitamin D pathways unbalanced in breast lesions. *BMC Cancer*, *10*. https://doi.org/10.1186/1471-2407-10-483

Lu, L., Sheng, H., Li, H., Gan, W., Liu, C., Zhu, J., … Lin, X. (2012). Associations between common variants in *GC* and *DHCR7/NADSYN1* and vitamin D concentration in Chinese Hans. *Human Genetics*, *131*(3), 505–512. https://doi.org/10.1007/s00439-011-1099-1

Machado, M. R. M., de Sousa Almeida-Filho, B., De Luca Vespoli, H., Schmitt, E. B., Nahas-Neto, J., & Nahas, E. A. P. (2019). Low pretreatment serum concentration of vitamin D

at breast cancer diagnosis in postmenopausal women. *Menopause (New York, N.Y.)*, *26*(3), 1–7. https://doi.org/10.1097/GME.0000000000001203

Mackawy, A. M. H., & Badawi, M. E. H. (2014). Association of vitamin D and vitamin D receptor gene polymorphisms with chronic inflammation, insulin resistance and metabolic syndrome components in type 2 diabetic Egyptian patients. *Meta Gene*, *2*(August 2016), 540–556. https://doi.org/10.1016/j.mgene.2014.07.002

Madden, J. M., Murphy, L., Zgaga, L., & Bennett, K. (2018). De novo vitamin D supplement use post-diagnosis is associated with breast cancer survival. *Breast Cancer Research and Treatment*, *172*(1), 179–190. https://doi.org/10.1007/s10549-018-4896-6

Maier, G. S., Horas, K., Kurth, A. A., Lazovic, D., Seeger, J. B., & Maus, U. (2015). Prevalence of vitamin D deficiency in patients with bone metastases and multiple myeloma. *Anticancer Research*, *35*(11), 6281–6285. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/26504063

Makarem, N., Chandran, U., Bandera, E. V., & Parekh, N. (2013). Dietary fat in breast cancer survival. *Annual Review of Nutrition*, *33*(1), 319–348. https://doi.org/10.1146/annurev-nutr-112912-095300

Makki, J. (2015). Diversity of breast carcinoma: Histological subtypes and clinical relevance. *Clinical Medicine Insights Pathology*, *8*(1), 23–31. https://doi.org/10.4137/CPath.s31563

Malhotra, G. K., Zhao, X., Band, H., & Band, V. (2010). Histological, molecular and functional subtypes of breast cancers. *Cancer Biology and Therapy*, *10*(10), 955–960. https://doi.org/10.4161/cbt.10.10.13879

Maneechay, W., Boonpipattanapong, T., Kanngurn, S., Puttawibul, P., Geater, S. L., & Sangkhathat, S. (2015). Single nucleotide polymorphisms in the *Gc* gene for vitamin D binding protein in common cancers in Thailand. *Asian Pacific Journal of Cancer Prevention*, *16*(8), 3339–3344. https://doi.org/10.7314/apjcp.2015.16.8.3339

Manson, J. E., Cook, N. R., Lee, I.-M., Christen, W., Bassuk, S. S., Mora, S., … Buring, J. E. (2019). Vitamin D supplements and prevention of cancer and cardiovascular disease. *New England Journal of Medicine*, *380*(1), 33–44. https://doi.org/10.1056/NEJMoa1809944

Mardis, E. R. (2017). DNA sequencing technologies: 2006-2016. *Nature Protocols*, *12*(2), 213–218. https://doi.org/10.1038/nprot.2016.182

Marozik, P. M., Tamulaitiene, M., Rudenka, E., Alekna, V., Mosse, I., Rudenka, A., … Kobets, K. (2018). Association of vitamin D receptor gene variation with osteoporosis risk in Belarusian and Lithuanian postmenopausal women. *Frontiers in Endocrinology*, *9*, 305. https://doi.org/10.3389/fendo.2018.00305

Mastracci, T. L., Tjan, S., Bane, A. L., O'Malley, F. P., & Andrulis, I. L. (2005). E-cadherin alterations in atypical lobular hyperplasia and lobular carcinoma in situ of the breast. *Modern Pathology*, *18*(6), 741–751. https://doi.org/10.1038/modpathol.3800362

Matthew Stenger. (2013). Effects of menarche and menopause on breast cancer risk: Meta-analysis findings - The ASCO Post. Retrieved June 5, 2019, from https://www.ascopost.com/issues/april-15-2013/effects-of-menarche-and-menopause-on-breast-cancer-risk-meta-analysis-findings/

McCart Reed, A. E., Kutasovic, J. R., Lakhani, S. R., & Simpson, P. T. (2015). Invasive lobular carcinoma of the breast: morphology, biomarkers and 'omics. *Breast Cancer Research*, *17*(12), 11. https://doi.org/10.1186/s13058-015-0519-x

McCullough, M. L., Stevens, V. L., Diver, W. R., Feigelson, H. S., Rodriguez, C., Bostick, R. M., … Calle, E. E. (2007). Vitamin D pathway gene polymorphisms, diet, and risk of postmenopausal breast cancer: a nested case-control study. *Breast Cancer Research : BCR*, *9*(1), R9. https://doi.org/10.1186/bcr1642

Mckee, A., Ribeiro, S. M. L., Malmstrom, T. K., , H.M. Perry III, D.K. Miller, Farr, S. S.,

Niehoff, M. L., & Albert, S. G. (2018). Screening for vitamin D deficiency in Black Americans: Comparison of total, free, bioavailable 25hydroxy vitamin D levels with parathyroid hormone levels and bone mineral density. *Journal of Nutrition Health and Aging*, *22*(9), 1045–1050. https://doi.org/10.1007/s12603-018-1080-y

Mehir un Nisa Iqbal and Taseer Ahmed Khan. (2017). Association between Vitamin D receptor (*Cdx2, Fok1, Bsm1, Apa1, Bgl1, Taq1, and Poly (A)*) gene polymorphism and breast cancer: A systematic review and meta-analysis. *Tumour Biology*, *39*(10), 1–9. https://doi.org/10.1177/1010428317731280

Mehrgou, A., & Akouchekian, M. (2016). The importance of *BRCA1* and *BRCA2* genes mutations in breast cancer development. *Medical Journal of the Islamic Republic of Iran*, *30*(15), 369. Retrieved from http://mjiri.iums.ac.ir

Mersch, J., Jackson, M. A., Park, M., Nebgen, D., Peterson, S. K., Singletary, C., … Litton, J. K. (2015). Cancers associated with *BRCA1* and *BRCA2* mutations other than breast and ovarian. *Cancer*, *121*(2), 269–275. https://doi.org/10.1002/cncr.29041

Miao, H., Verkooijen, H. M., Chia, K. S., Bouchardy, C., Pukkala, E., Larønningen, S., … Hartman, M. (2011). Incidence and outcome of male breast cancer: An international population-based study. *Journal of Clinical Oncology*, *29*(33), 4381–4386. https://doi.org/10.1200/JCO.2011.36.8902

Mo, D., Hsieh, P., Yu, H., Zhou, L., Gong, J., Xu, L., … Deng, Q. (2017). The relationship between osteoporosis and body composition in pre- and postmenopausal women from different ethnic groups in China. *Ethnicity & Health*, *22*(3), 295–310. https://doi.org/10.1080/13557858.2016.1244758

Moore, L. D., Le, T., & Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology*, *38*(1), 23–38. https://doi.org/10.1038/npp.2012.112

Moukayed, M., & Grant, W. B. (2013). Molecular link between vitamin D and cancer

prevention. *Nutrients*, *5*(10), 3993–4021. https://doi.org/10.3390/nu5103993

Narvaez, C. J., Matthews, D., Laporta, E., Simmons, K. M., Beaudin, S., Welsh, J., & Carlberg, C. (2014). The impact of vitamin D in breast cancer: genomics, pathways, metabolism. *Frontiers in Physiology*, *5*, 1–10. https://doi.org/10.3389/fphys.2014.00213

Newman, L. A. (2015). Disparities in breast cancer and African ancestry: A global perspective. *Breast Journal*, *21*(2), 133–139. https://doi.org/10.1111/tbj.12369

Niehoff, N. M., Nichols, H. B., Zhao, S., White, A. J., & Sandler, D. P. (2019). Adult physical activity and breast cancer risk in women with a family history of breast cancer. *Cancer Epidemiology Biomarkers & Prevention*, *28*(1), 51–58. https://doi.org/10.1158/1055-9965.EPI-18-0674

Nissen, J., Rasmussen, L. B., Ravn-Haren, G., Wreford Andersen, E., Hansen, B., Andersen, R., … Vogel, U. (2014). Common variants in *CYP2R1* and *GC* genes predict vitamin D concentrations in healthy Danish children. *PLoS ONE*, *9*(2), e89907. https://doi.org/10.1371/journal.pone.0089907

Norval, M., Coussens, A. K., Wilkinson, R. J., Bornman, L., Lucas, R. M., & Wright, C. Y. (2016). Vitamin D status and its consequences for health in South Africa. *International Journal of Environmental Research and Public Health*, *13*(10), 1-22. https://doi.org/10.3390/ijerph13101019

Nykamp, K., Anderson, M., Powers, M., Garcia, J., Herrera, B., Ho, Y. Y., … Topper, S. (2017). Sherloc: A comprehensive refinement of the ACMG-AMP variant classification criteria. *Genetics in Medicine*, *19*(10), 1105–1117. https://doi.org/10.1038/gim.2017.37

O'brien, K. M., Sandler, D. P., Kinyamu, H. K., Taylor, J. A., & Weinberg, C. R. (2017). Single-nucleotide polymorphisms in vitamin D-related genes may modify vitamin D-breast cancer associations. *Cancer Epidemiology Biomarkers & Prevention*, *26*(12), 1761–1772. https://doi.org/10.1158/1055-9965.EPI-17-0250

O'Brien, K. M., Sandler, D. P., Shi, M., Harmon, Q. E., Taylor, J. A., & Weinberg, C. R. (2018). Genome-wide association study of serum 25-hydroxyvitamin D in US women. *Frontiers in Genetics*, *9*(MAR), 1–11. https://doi.org/10.3389/fgene.2018.00067

Ogunkorode, A., Holtslander, L., Anonson, J., & Maree, J. (2017). An integrative review of the literature on the determinants of health outcomes of women living with breast cancer in Canada and Nigeria from 1990 to 2014: A comparative study. *International Journal of Africa Nursing Sciences*, *6*, 52–73. https://doi.org/10.1016/j.ijans.2017.04.002

Ordóñez-Mena, J. M., Schöttker, B., Fedirko, V., Jenab, M., Olsen, A., Halkjær, J., … Brenner, H. (2016). Pre-diagnostic vitamin D concentrations and cancer risks in older individuals: an analysis of cohorts participating in the CHANCES consortium. *European Journal of Epidemiology*, *31*(3), 311–323. https://doi.org/10.1007/s10654-015-0040-7

Palomer, X., González-Clemente, J. M., Blanco-Vaca, F., & Mauricio, D. (2008). Role of vitamin D in the pathogenesis of type 2 diabetes mellitus. *Diabetes, Obesity and Metabolism*, *10*(3), 185–197. https://doi.org/10.1111/j.1463-1326.2007.00710.x

Park, S., Dong, Lee, H., Jeon, J. Y., Ryu, J., Kim, S., … Park, B.-W. (2015). Serum 25-hydroxyvitamin D deficiency and increased risk of breast cancer among Korean women: a case-control study. *Breast Cancer Research and Treatment*, *152*, 147–154. https://doi.org/10.1007/s10549-015-3433-0

Parra, E. J. (2007). Human pigmentation variation: Evolution, genetic basis, and implications for public health. *American Journal of Physical Anthropology*, *134*(s45), 85–105. https://doi.org/10.1002/ajpa.20727

Patel, A., Belykh, E., Miller, E. J., George, L. L., Martirosyan, N. L., Byvaltsev, V. A., & Preul, M. C. (2018). MinION rapid sequencing: Review of potential applications in neurosurgery. *Surgical Neurology International*, *9*, 157-165.

https://doi.org/10.4103/sni.sni_55_18

Pedrera-Canal, M., Moran, J. M., Vera, V., Roncero-Martin, R., Lavado-Garcia, J. M., Aliaga, I., & Pedrera-Zamorano, J. D. (2015). Common allelic variants of the vitamin receptor D gene *rs7975232* (ApaI) do not influence bone mineral density figures in postmenopausal osteoporotic women. *International Journal of Clinical and Experimental Medicine*, *8*(5), 8173–8177. http://www.ncbi.nlm.nih.gov/pubmed/26221388

Peñ, C., García, J. M., Silva, J., García, V., Rodríguez, R., Alonso, I., … Lix Bonilla, F. (2005). E-cadherin and vitamin D receptor regulation by *SNAIL* and *ZEB1* in colon cancer: clinicopathological correlations. *Human Molecular Genetics*, *14*(22), 3361–3370. https://doi.org/10.1093/hmg/ddi366

Perou, C. M., Sørlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. a, … Botstein, D. (2000). Molecular portraits of human breast tumours. *Nature*, *406*(6797), 747–752. https://doi.org/10.1038/35021093

Pervin, S., Hewison, M., Braga, M., Tran, L., Chun, R., Karam, A., … Singh, R. (2013). Down-regulation of vitamin D receptor in mammospheres: Implications for vitamin D resistance in breast cancer and potential for combination therapy. *PLoS ONE*, *8*(1), e53287. https://doi.org/10.1371/journal.pone.0053287

Pickholtz, I., Saadyan, S., Keshet, G. I., Wang, V. S., Cohen, R., Bouwman, P., … Yarden, R. I. (2014). Cooperation between *BRCA1* and vitamin D is critical for histone acetylation of the p21waf1 promoter and growth inhibition of breast cancer cells and cancer stem-like cells. *Oncotarget*, *5*(23), 11827–11846. https://doi.org/10.18632/oncotarget.2582

Pilz, S., Kienreich, K., Rutters, F., De Jongh, R., Van Ballegooijen, A. J., Grübler, M., … Dekker, J. M. (2013). Role of vitamin D in the development of insulin resistance and

type 2 diabetes. *Current Diabetes Reports*, *13*(2), 261–270. https://doi.org/10.1007/s11892-012-0358-4

Pittas, A. G., Dawson-Hughes, B., Sheehan, P., Ware, J. H., Knowler, W. C., Aroda, V. R., … Staten, M. (2019). Vitamin D supplementation and prevention of type 2 diabetes. *New England Journal of Medicine*, *381*(6), 520–530. https://doi.org/10.1056/NEJMoa1900906

Plesivkova, D., Richards, R., & Harbison, S. (2019). A review of the potential of the MinION™ single-molecule sequencing system for forensic applications. *Wiley Interdisciplinary Reviews: Forensic Science*, *1*(1), e1323. https://doi.org/10.1002/wfs2.1323

Plon, S. E., Eccles, D. M., Easton, D., Foulkes, W. D., Genuardi, M., Greenblatt, M. S., … IARC unclassified genetic variants working group. (2008). Sequence variant classification and reporting: Recommendations for improving the interpretation of cancer susceptibility genetic test results. *Human Mutation*, *29*(11), 1282–1291. https://doi.org/10.1002/humu.20880

Pohl, H., Kotze, M. J., Grant, K. A., van der Merwe, L., Pienaar, F. M., Apffelstaedt, J. P., & Myburgh, E. J. (2016). Impact of mammaprint on clinical decision-making in South African patients with early-stage breast cancer. *The Breast Journal*, 22(4), 442-446. https://doi.org/10.1111/tbj.12605

Praetorius, C., Grill, C., Stacey, S. N., Metcalf, A. M., Gorkin, D. U., Robinson, K. C., … Steingrimsson, E. (2013). A polymorphism in IRF4 affects human pigmentation through a tyrosinase-dependent MITF/TFAP2A Pathway. *Cell*, *155*(5), 1022–1033. https://doi.org/10.1016/J.CELL.2013.10.022

Prieto-Alhambra, D., Kassim Javaid, M., Servitja, S., Arden, N. K., Martinez-García, M., Diez-Perez, A., … Nogues, X. (2011). Vitamin D threshold to prevent aromatase

inhibitor-induced arthralgia: A prospective cohort study. *Breast Cancer Res Treat*, *125*, 869–878. https://doi.org/10.1007/s10549-010-1075-9

Quail, M. A., Smith, M., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., … Gu, Y. (2012). A tale of three next generation sequencing platforms: Comparison of ion torrent, pacific biosciences and illumina miseq sequencers. *Genomics*, *13*, 341. https://doi.org/10.1186/1471-2164-13-341

Rakha, E. A., Reis-Filho, J. S., & Ellis, I. O. (2010). Combinatorial biomarker expression in breast cancer. *Breast Cancer Research and Treatment*, 120(2), 293-308. https://doi.org/10.1007/s10549-010-0746-x

Rakha EA and Ellis IO. (2010). Lobular breast carcinoma and its variants. *Seminars in Diagnostic Pathology*, *27*(1), 49–61. https://doi.org/10.1053/J.SEMDP.2009.12.009

Rang, F. J., Kloosterman, W. P., & de Ridder, J. (2018). From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. *Genome Biology*, *19*(1), 90. https://doi.org/10.1186/s13059-018-1462-9

Rashid, M. U., Muzaffar, M., Khan, F. A., Kabisch, M., Muhammad, N., Faiz, S., … Hamann, U. (2015). Association between the *BsmI* polymorphism in the vitamin D receptor gene and breast cancer risk: Results from a Pakistani Case-Control Study. *PLoS ONE*, *10*(10), e0141562. https://doi.org/10.1371/journal.pone.0141562

Redaniel, M. T., Gardner, M. P., Martin, R. M., & Jeffreys, M. (2014). The association of vitamin D supplementation with the risk of cancer in postmenopausal women. *Cancer Causes Control*, *25*, 267–271. https://doi.org/10.1007/s10552-013-0328-4

Reid, D. M., Doughty, J., Eastell, R., Heys, S. D., Howell, A., Mccloskey, E. V, … Coleman, R. E. (2008). Guidance for the management of breast cancer treatment-induced bone loss: A consensus position statement from a UK Expert Group q. *Cancer Treatment Reviews*, *34*, S3–S18. https://doi.org/10.1016/j.ctrv.2008.03.007

Reimers, L. L., Crew, K. D., Patrick, Bradshaw, T., Santella, R. M., Steck, S. E., … Gammon, M. D. (2015). Vitamin D-related gene polymorphisms, plasma 25-hydroxyvitamin D, and breast cancer risk. *Cancer Causes Control*, *26*, 187–203. https://doi.org/10.1007/s10552-014-0497-9

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., … ACMG laboratory quality assurance committee. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American college of medical genetics and genomics and the association for molecular pathology. *Genetics in Medicine*, *17*(5), 405–423. https://doi.org/10.1038/gim.2015.30

Rizzoli R, Body JJ, DeCensi A, Reginster JY, Piscitelli P, B. M. (2012). Guidance for the prevention of bone loss and fractures in postmenopausal women treated with aromatase inhibitors for breast cancer: an ESCEO position paper. *Osteoporosis International*, *23*(11), 2567–2576. https://doi.org/10.1007/s00198-011-1870-0

Rochel, N., & Molnár, F. (2017). Structural aspects of vitamin D endocrinology. *Molecular and Cellular Endocrinology*, *453*, 22–35. https://doi.org/10.1016/j.mce.2017.02.046

Rosai, J., & Ackerman, L. V. (2011). *Rosai and Ackerman's surgical pathology*. Mosby.

Rossberg, W., Saternus, R., Wagenpfeil, S., Kleber, M., März, W., Reichrath, S., … Reichrath, J. (2016). Human pigmentation, cutaneous vitamin d synthesis and evolution: variants of genes (SNPs) involved in skin pigmentation are associated with 25(OH)D serum concentration. *Anticancer Research*, *36*(3), 1429–1437. http://www.ncbi.nlm.nih.gov/pubmed/26977047

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors (DNA polymerase/nucleotide sequences/bacteriophage 4X174). *Proceedings of the National Academy of Sciences of the United States of America*, *74*(12), 5463–5467. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC431765/pdf/pnas00043-0271.pdf

Santos, B. R., Lecke, S. B., & Spritzer, P. M. (2017). Genetic variant in vitamin D-binding protein is associated with metabolic syndrome and lower 25-hydroxyvitamin D levels in polycystic ovary syndrome: A cross-sectional study. *PloS One*, *12*(3), e0173695. https://doi.org/10.1371/journal.pone.0173695

Saternus, R., Pilz, S., Gräber, S., Kleber, M., März, W., Vogt, T., & Rg Reichrath, J. (2015). A closer look at evolution: Variants (SNPs) of genes involved in skin pigmentation, including *EXOC2, TYR, TYRP1*, and *DCT,* are associated with 25(OH)D serum concentration. *Endocrinology*, *156*, 39–47. https://doi.org/10.1210/en.2014-1238

Schoeman, M., Apffelstaedt, J. P., Baatjes, K., & Urban, M. (2013). Implementation of a breast cancer genetic service in South Africa - lessons learned. *South African Medical Journal*, *103*(8), 529–533. https://doi.org/10.7196/SAMJ.6814

Schübeler, D. (2015). Function and information content of DNA methylation. *Nature*, *517*(7534), 321–326. https://doi.org/10.1038/nature14192

Servitja, S., Nogués, X., Prieto-Alhambra, D., Martínez-García, M., Garrigós, L., Peña, M. J., … Tusquets, I. (2012). Bone health in a prospective cohort of postmenopausal women receiving aromatase inhibitors for early breast cancer. *Breast (Edinburgh, Scotland)*, *21*(1), 95–101. https://doi.org/10.1016/j.breast.2011.09.001

Shahabi, A., Alipour, M., Safiri, H., Tavakol, P., Alizadeh, M., Hashemi, & S. M., … Halimi, M. (2018). Vitamin D receptor gene polymorphism: Association with susceptibility to early-onset breast cancer in Iranian, *BRCA1/2*-mutation carrier and non-carrier patients. *Pathol. Oncol. Res.*, *24*(3), 601–607. https://doi.org/10.1007/s12253-017-0281-8

Shi, J., Grundy, A., Richardson, H., Burstyn, I., Schuetz, J. M., Lohrisch, C. A., … Ca, A. (2016). Genetic variation in vitamin D-related genes and risk of breast cancer among women of European and East Asian descent. *Tumor Biology*, *37*, 6379–6387. https://doi.org/10.1007/s13277-015-4417-8

Shi, L., Nechuta, S., Gao, Y. T., Zheng, Y., Dorjgochoo, T., Wu, J., … Shu, X. O. (2014). Correlates of 25-hydroxyvitamin D among Chinese breast cancer patients. *PLoS ONE*, *9*(1), 14–16. https://doi.org/10.1371/journal.pone.0086467

Shiovitz, S., & Korde, L. A. (2015). Genetics of breast cancer: A topic in evolution. *Annals of Oncology*, *26*(7), 1291–1299. https://doi.org/10.1093/annonc/mdv022

Signorello, L. B., Shi, J., Cai, Q., Zheng, W., Williams, S. M., Long, J., … Blot, W. J. (2011). Common variation in vitamin D pathway genes predicts circulating 25-hydroxyvitamin D levels among African Americans. *PLoS ONE*, *6*(12), 13–19. https://doi.org/10.1371/journal.pone.0028623

Sillanpaa, P. (2007). Polymorphic low penetrance genes and breast cancer - the role of genes involved in metabolism of xenobiotics, estrogens and reactive oxygen species. *People and Work Research Reports 78*, 1–123.

Singleton, A. B. (2011). Exome sequencing: A transformative technology. *Lancet Neurology*, *10*, 942–946. http://www.ncbi.nlm.nih.

Sinotte, M., Rousseau, F., Ayotte, P., Dewailly, E., Diorio, C., Giguère, Y., … Brisson, J. (2008). Vitamin D receptor polymorphisms (*FokI, BsmI*) and breast cancer risk: Association replication in two case-control studies within French Canadian population. *Endocrine-Related Cancer*, *15*(4), 975–983. https://doi.org/10.1677/ERC-08-0056

Smith, K. R., Hanson, H. A., & Hollingshaus, M. S. (2013). *BRCA1* and *BRCA2* mutations and female fertility. *Current Opinion in Obstetrics & Gynecology*, *25*(3), 207–213. https://doi.org/10.1097/GCO.0b013e32835f1731

Sperati, F., Vici, P., Maugeri-Saccà, M., Stranges, S., & Santesso, N. (2013). Vitamin D supplementation and breast cancer prevention: A systematic review and meta-analysis of randomized clinical trials. *PLoS ONE*, *8*(7), e69269. https://doi.org/10.1371/journal.pone.0069269

Spiro, A., & Buttriss, J. L. (2014). Vitamin D: An overview of vitamin D status and intake in Europe. *Nutrition Bulletin*, *39*(4), 322–350. https://doi.org/10.1111/nbu.12108

Sporle, A., Ellison-Loschmann, L., Koea, J., Guilford, P., Day, R., Cheng, S., … Pearce, N. (2018). Germline *CDH1* mutations are a significant contributor to the high frequency of early-onset diffuse gastric cancer cases in New Zealand Māori. *Familial Cancer*, *18*(1), 83–90. https://doi.org/10.1007/s10689-018-0080-8

Strauss, J. F. (2014). The synthesis and metabolism of steroid hormones. *Yen & Jaffe's Reproductive Endocrinology* (Seventh Ed). Elsevier. https://doi.org/10.1016/B978-1-4557-2758-2.00004-4

Stray-Pedersen, A., Sorte, H. S., Samarakoon, P., Gambin, T., Chinn, I. K., Coban Akdemir, Z. H., … Lupski, J. R. (2017). Primary immunodeficiency diseases: Genomic approaches delineate heterogeneous Mendelian disorders. *Journal of Allergy and Clinical Immunology*, *139*(1), 232–245. https://doi.org/10.1016/J.JACI.2016.05.042

Sun, Y. S., Zhao, Z., Yang, Z. N., Xu, F., Lu, H. J., Zhu, Z. Y., … Zhu, H. P. (2017). Risk factors and preventions of breast cancer. *International Journal of Biological Sciences*, *13*(11), 1387–1397. https://doi.org/10.7150/ijbs.21635

Surakasula, A., Nagarjunapu, G. C., & Raghavaiah, K. V. (2014). A comparative study of pre- and post-menopausal breast cancer: Risk factors, presentation, characteristics and management. *Journal of Research in Pharmacy Practice*, *3*(1), 12–18. https://doi.org/10.4103/2279-042X.132704

Swami, Srilatha, Aruna V. Krishnan, Jennifer Y. Wang, et al. (2011). Inhibitory effects of calcitriol on the growth of MCF-7 breast cancer xenografts in nude mice: Selective modulation of aromatase expression in vivo. *Hormones and Cancer*, *2*(3), 190–202. https://doi.org/10.1007/s12672-011-0073-7

Swami, S., Sarabia, S. F., Diamandis, A., Mistry, J., Khosravi, J., & Feldman, D. (2001). A

new enzyme-linked immunosorbant assay for the measurement of human vitamin D receptor. *Bone*, *28*(3), 319–326. https://doi.org/10.1016/S8756-3282(00)00432-4

Swami, Srilatha, Krishnan, A. V., & Feldman, D. (2011). Vitamin D metabolism and action in the prostate: Implications for health and disease. *Molecular and Cellular Endocrinology*, *347*(1–2), 61–69. https://doi.org/10.1016/j.mce.2011.05.010

Tagliabue, E., Raimondi, S., & Gandini, S. (2015). Meta-analysis of vitamin D-binding protein and cancer risk. *Cancer Epidemiology Biomarkers & Prevention*, *24*(11), 1758–1765. https://doi.org/10.1158/1055-9965.EPI-15-0262

Tang, C., Chen, N., Wu, M., Yuan, H., & Du, Y. (2009). *Fok1* polymorphism of vitamin D receptor gene contributes to breast cancer susceptibility: A meta-analysis. *Breast Cancer Research and Treatment*, *117*(2), 391–399. https://doi.org/10.1007/s10549-008-0262-4

Tavassoli Fattaneh A and Devilee Peter. (2003). *WHO Classification Pathology and Genetics of Tumours of the Breast and Female Genital Organs.* (D. P. Tavassoli FA, Ed.), *IARC press* (3rd editio, Vol. 193). Lyon: WHO press. Retrieved from https://breast-cancer-research.biomedcentral.com/track/pdf/10.1186/bcr788

Tavera-Mendoza, L. E., Westerling, T., Libby, E., Marusyk, A., Cato, L., Cassani, R., … Brown, M. (2017). Vitamin D receptor regulates autophagy in the normal mammary gland and in luminal breast cancer cells. *Proceedings of the National Academy of Sciences*, 201615015. https://doi.org/10.1073/pnas.1615015114

Taylor, A., Brady, A. F., Frayling, I. M., Hanson, H., Tischkowitz, M., Turnbull, C., & Side, L. (2018). Consensus for genes to be included on cancer panel tests offered by UK genetics services: guidelines of the UK Cancer Genetics Group. *Journal of Medical Genetics*, *55*(6), 372–377. https://doi.org/10.1136/jmedgenet-2017-105188

Thanasitthichai, S., Chaiwerawattana, A., & Prasitthipayong, A. (2015). Association of

vitamin D level with clinicopathological features in breast cancer. *Asian Pacific Journal of Cancer Prevention*, 16(12), 4881–4883.

The Cancer Genome Atlas Network. (2012). Comprehensive molecular portraits of human breast tumours. *Nature*, 490(7418):61-70. https://doi.org/10.1038/nature11412

Thomas, R. J., Kenfield, S. A., & Jimenez, A. (2017). Exercise-induced biochemical changes and their potential influence on cancer: A scientific review. *British Journal of Sports Medicine*, *51*(8), 640–644. https://doi.org/10.1136/bjsports-2016-096343

Thorne, J., & Campbell, M. J. (2008). The vitamin D receptor in cancer. *The Proceedings of the Nutrition Society*, *67*(2), 115–127. https://doi.org/10.1017/S0029665108006964

Tommie, J. L., Pinney, S. M., & Nommsen-Rivers, L. A. (2018). Serum vitamin D status and breast cancer risk by receptor status: A systematic review. *Nutrition and Cancer*, *70*(5), 804–820. https://doi.org/10.1080/01635581.2018.1470653

Tung, N., Wang, Y., Collins, L. C., Kaplan, J., Li, H., Gelman, R., … Schnitt, S. J. (2010). Estrogen receptor positive breast cancers in *BRCA1* mutation carriers: Clinical risk factors and pathologic features. *Breast Cancer Research*, *12*(1), R12. https://doi.org/10.1186/bcr2478

Turashvili, G., & Brogi, E. (2017). Tumor heterogeneity in breast cancer. *Tumor Heterogeneity in Breast Cancer. Frontiers in Medicine*, *4*, 1-11. https://doi.org/10.3389/fmed.2017.00227

Uitterlinden, A. G., Fang, Y., Van Meurs, J. B. J., Pols, H. A. P., & Van Leeuwen, J. P. T. M. (2004). Genetics and biology of vitamin D receptor polymorphisms. *Gene*, *338*(2), 143–156. https://doi.org/10.1016/j.gene.2004.05.014

un Nisa Iqbal, Mehir, Amir Maqbool, S., & Ahmed Khan, T. (2018). Association of low penetrance vitamin D receptor Tru9I (*rs757343*) gene polymorphism with risk of

premenopausal breast cancer. *Journal of International Medical Research*, *46*(5), 1801–1814. https://doi.org/10.1177/0300060518761304

un Nisa Iqbal, Mehr, Ahmed Khan, T., & Amir Maqbool, S. (2015). Vitamin D receptor *Cdx-2* polymorphism and premenopausal breast cancer risk in Southern Pakistani patients. *PLoS ONE*, *10*(3), e0122657. https://doi.org/10.1371/journal.pone.0122657

Usman Rashid, M., Muzaffar, M., Ali Khan, F., Kabisch, M., Muhammad, N., Faiz, S., … Hamann, U. (2015). Association between the *BsmI* polymorphism in the vitamin D receptor gene and breast cancer risk: Results from a Pakistani case-control study. *PLoS ONE*, *10*(10), e0141562. https://doi.org/10.1371/journal.pone.0141562

Vallejo, M. S., Blümel, J. E., Lavín, P., Torres, C., Araos, A., & Sciaraffia, C. (2018). Older women do not have seasonal variations of vitamin D levels. *Menopause*, *25*(8), 912–917. https://doi.org/10.1097/GME.0000000000001103

van der Merwe, N., Peeters, A. V., Pienaar, F. M., Bezuidenhout, J., Van Rensburg, S. J., & Kotze, M. J. (2017). Exome sequencing in a family with luminal-type breast cancer underpinned by variation in the methylation pathway. *International Journal of Molecular Sciences*, *18*(2), 467-483. https://doi.org/10.3390/ijms18020467

van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., & Thermes, C. (2018). The third revolution in sequencing technology. *Trends in Genetics*, *34*(9), 666–681. https://doi.org/10.1016/j.tig.2018.05.008

Vaughan-Shaw, P. G., O'Sullivan, F., Farrington, S. M., Theodoratou, E., Campbell, H., Dunlop, M. G., & Zgaga, L. (2017). The impact of Vitamin D pathway genetic variation and circulating 25-hydroxyVitamin D on cancer outcome: Systematic review and meta-Analysis. *British Journal of Cancer*, *116*(8), 1095–1110. https://doi.org/10.1038/bjc.2017.44

Vimaleswaran, K. S., Cavadino, A., Berry, D. J., Power, C., Hyppnen, E., Jorde, R., …

Hyppnen, E. (2014). Association of vitamin D status with arterial blood pressure and hypertension risk: A mendelian randomisation study. *The Lancet Diabetes and Endocrinology*, *2*(9), 719–729. https://doi.org/10.1016/S2213-8587(14)70113-5

Visser, J., Knight, K., Philips, L., Visser, W., Wallace, M., Nel, D., & Blaauw, R. (2019). Determinants of serum 25-hydroxyvitamin D levels in healthy young adults living in the Western Cape, South Africa. *South African Family Practice*, *61*(4), 150–158. https://doi.org/10.1080/20786190.2019.1621047

Wacker, M., & Holick, M. F. (2013). Sunlight and vitamin D: A global perspective for health. *Dermato-Endocrinology*, *5*(1), 51–108. https://doi.org/10.4161/derm.24476

Wang, H., Naghavi, M., Allen, C., Barber, R. M., Carter, A., Casey, D. C., … Zuhlke, L. J. (2016). Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: A systematic analysis for the global burden of disease study 2015. *The Lancet*, *388*(10053), 1459–1544. https://doi.org/10.1016/S0140-6736(16)31012-1

Wang, Y., Zhu, J., & DeLuca, H. F. (2012). Where is the vitamin D receptor? *Archives of Biochemistry and Biophysics*, *523*(1), 123–133. https://doi.org/10.1016/j.abb.2012.04.001

Weaver, S. A. L. and C. M. (2013). Global nutrition research: Nutrition and breast cancer prevention as a model. *Nutrition reviews*, *71*(11), 742–752. https://doi.org/10.1111/nure.12075.Global

Weber, B. L., & Nathanson, K. L. (2000). Low penetrance genes associated with increased risk for breast cancer. *European Journal of Cancer*, *36*(10), 1193–1199. https://doi.org/10.1016/S0959-8049(00)00082-4

Wilson, R. T., Roff, A. N., Dai, P. J., Fortugno, T., Douds, J., Chen, G., … Cheng, K. C. (2011). Genetic ancestry, skin reflectance and pigmentation genotypes in association

with serum vitamin D metabolite balance. *Hormone Molecular Biol ology and Clinical Investigation*, *7*(1), 279–293. https://doi.org/10.1515/HMBCI.2011.021

World Cancer Research Fund/American Institute for Cancer Research. (2018). Diet, nutrition, physical activity and breast cancer. *Continuous Update Project Expert Report 2018*, 1-50. https://doi.org/10.1007/s12082-007-0105-4

Yanase, T., Suzuki, S., Goto, K., Nomura, M., Okabe, T., Takayanagi, R., & Nawata, H. (2003). Aromatase in bone: Roles of vitamin D_3 and androgens. *Journal of Steroid Biochemistry and Molecular Biology*, *86*(3–5), 393–397. https://doi.org/10.1016/S0960-0760(03)00349-2

Yao, S., Hong, C.-C., Bandera, E. V, Zhu, Q., Liu, S., David Cheng, T.-Y., … Ambrosone, C. B. (2017). Demographic, lifestyle, and genetic determinants of circulating concentrations of 25-hydroxyvitamin D and vitamin D-binding protein in African American and European American women. *American Journal of Clinical Nutrition*, *105*, 1362–1371. https://doi.org/10.3945/ajcn.116.143248

Yao, S., Sucheston, L. E., Millen, A. E., Johnson, C. S., Trump, D. L., Nesline, M. K., … Ambrosone, C. B. (2011). Pretreatment serum concentrations of 25-hydroxyvitamin D and breast cancer prognostic characteristics: A case-control and a case-series study. *PLoS ONE*, *6*(2), e17251. https://doi.org/10.1371/journal.pone.0017251

Ye, Z., Sharp, S. J., Burgess, S., Scott, R. A., Imamura, F., Langenberg, C., … Forouhi, N. G. (2015). Association between circulating 25-hydroxyvitamin D and incident type 2 diabetes: A mendelian randomisation study. *The Lancet Diabetes and Endocrinology*, *3*(1), 35–42. https://doi.org/10.1016/S2213-8587(14)70184-6

Yousefzadeh, P., Shapses, S. A., & Wang, X. (2014). Vitamin D binding protein impact on 25-hydroxyvitamin D levels under different physiologic and pathologic conditions. *International journal of endocrinology*, *2014*, 1-6. https://doi.org/10.1155/2014/981581

Zhang, F., & Finkelstein, J. (2019). Inconsistency in race and ethnic classification in pharmacogenetics studies and its potential clinical implications. *Pharmacogenomics and Personalized Medicine*,*12*, 107–123. https://doi.org/10.2147/PGPM.S207449

Zhang, L., Yin, X., Wang, J., Xu, D., Wang, Y., Yang, J., … Yan, C. (2018). Associations between *VDR* gene polymorphisms and osteoporosis risk and bone mineral density in postmenopausal women: A systematic review and meta-analysis. *Scientific Reports*, *8*(1), 981. https://doi.org/10.1038/s41598-017-18670-7

Zhang, Z., He, J.-W., Fu, W.-Z., Zhang, C.-Q., & Zhang, Z.-L. (2013). An analysis of the association between the vitamin D pathway and serum 25-hydroxyvitamin D levels in a healthy Chinese population. *Journal of Bone and Mineral Research*, *28*(8), 1784–1792. https://doi.org/10.1002/jbmr.1926

Zhou, Z.-C., Wang, J., Cai, Z.-H., Zhang, Q.-H., Cai, Z.-X., & Wu, J.-H. (2013). Association between vitamin D receptor gene *Cdx2* polymorphism and breast cancer susceptibility. *Tumor Biology*, *34*(6), 3437–3441. https://doi.org/10.1007/s13277-013-0919-4

Zhu, M., Tan, Z., Luo, Z., Hu, H., Wu, T., Fang, S., … Lu, Z. (2019). Association of the vitamin D metabolism gene *GC* and *CYP27B1* polymorphisms with cancer susceptibility: A meta-analysis and trial sequential analysis. *Bioscience Reports*, 39(9), 1–13. https://doi.org/10.1042/BSR20190368

Zhuang, J. C., Huang, Z. Y., Zhao, G. X., Yu, H., Li, Z. X., & Wu, Z. Y. (2015). Variants of *CYP27B1* are associated with both multiple sclerosis and neuromyelitis optica patients in Han Chinese population. *Gene*, *557*(2), 236–239. https://doi.org/10.1016/j.gene.2014.12.045

APPENDICES

APPENDIX I

UNIVERSITEIT
STELLENBOSCH
UNIVERSITY

Progress Report Approval Letter

11/12/2018

Project Reference #: 1352

Ethics Reference #: S17/10/219

Title: Investigation of the role of vitamin D metabolism in South African breast cancer patients using a pathology-supported genetic testing platform

Dear Mr. Abisola Okunola,

Your request for extension/annual renewal of ethics approval dated 05/12/2018 15:35 refers.

The Health Research Ethics Committee reviewed and approved the annual progress report you submitted through an expedited review process.

The approval of this project is extended for a further year.

Approval date: 11 December 2018

Expiry date: 10 December 2019

Kindly be reminded to submit progress reports two (2) months before expiry date.

Where to submit any documentation

Kindly note that the HREC uses an electronic ethics review management system, *Infonetica*, to manage ethics applications and ethics review process. To submit any documentation to HREC, please click on the following link: https://applyethics.sun.ac.za.

Please remember to use your **Project ID** [1352] and Ethics Reference Number S17/10/219 on any documents or correspondence with the HREC concerning your research protocol.

National Health Research Ethics Council (NHREC) Registration Numbers: REC-130408-012 for HREC1 and REC-230208-010 for HREC2

Federal Wide Assurance Number: 00001372

Institutional Review Board (IRB) Number: IRB0005240 for HREC1

Institutional Review Board (IRB) Number: IRB0005239 for HREC2

The Health Research Ethics Committee complies with the SA National Health Act No. 61 of 2003 as it pertains to health research and the United States Code of Federal Regulations Title 45 Part 46. This committee abides by the ethical norms and principles for research, established by the Declaration of Helsinki and the South African Medical Research Council Guidelines as well as the Guidelines for Ethical Research: Principles, Structures and Processes 2015 (Department of Health).

Yours sincerely,

Mrs. Ashleen Fortuin

Health Research Ethics Committee 2 (HREC2).

UNIVERSITEIT
STELLENBOSCH
UNIVERSITY

Amendment Approval Letter

25/01/2019

Project ID: 1352

Ethics Reference #: S17/10/219

Title: Investigation of the role of vitamin D metabolism in South African breast cancer patients using a pathology-supported genetic testing platform

Dear Mr. Abisola Okunola,

Your amendment request dated 20 December 2018 refers.

The Health Research Ethics Committee (HREC) reviewed and approved the amended documentation through an expedited review process.

The following amendments were reviewed and approved:

1. Inclusion of all patients in Dr Baatjes Database in this study
2. Addition of Dr Armand Peeters as a Sub-Investigator to the study team.

Where to submit any documentation

Kindly note that the HREC uses an electronic ethics review management system, *Infonetica*, to manage ethics applications and ethics review process. To submit any documentation to HREC, please click on the following link: https://applyethics.sun.ac.za.

Please remember to use your **Project ID [1352]** and ethics reference number **[S17/10/219]** on any documents or correspondence with the HREC concerning your research protocol.

National Health Research Ethics Council (NHREC) Registration Numbers: REC-130408-012 for HREC1 and REC-230208-010 for HREC2

Federal Wide Assurance Number: 00001372

Institutional Review Board (IRB) Number: IRB0005240 for HREC1 Institutional Review Board (IRB) Number: IRB0005239 for HREC2
The Health Research Ethics Committee complies with the SA National Health Act No. 61 of 2003 as it pertains to health research and the United States Code of Federal Regulations Title 45 Part 46. This committee abides by the ethical norms and principles for research, established by the Declaration of Helsinki and the South African Medical Research Council Guidelines as well as the Guidelines for Ethical Research: Principles, Structures and Processes 2015 (Department of Health).

Yours sincerely,

Francis Masiye,

HREC Coordinator,

Health Research Ethics Committee 2 (HREC2).

APPENDIX II

STANDARD OPERATING PROCEDURE FOR DNA EXTRACTION

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. Five mls 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. Six mls 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 transferred 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. Five mls 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. Six hundred 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.

SANGER SEQUENCING STANDARD OPERATING PROCEDURES

Post PCR Clean-up: Post PCR clean-ups are performed using the Nucleofast 96 well post PCR clean-up plate from Macherey Nagel. The protocols supplied by the manufacturer are implemented on a Tecan EVO150 robotic workstation.

DNA Sequencing reactions: DNA sequencing is done using the BigDye Terminator V3.1 sequencing kit (Applied Biosystems) using the manufacturers protocol with slight modifications.

Post sequencing clean-up: After cycle sequencing the products are treated with SDS before it is transferred onto Sephadex columns (Princeton Scientific). The protocols supplied by the manufacturer are implemented on a Tecan EVO150 robotic workstation.

DNA sequencing electrophoresis: Prior to electrophoresis the samples are denatured for 2 minutes at 95°C and then cooled to 4°C. Electrophoresis is performed on an ABI3730xl using a 50cm capillary array and POP7 (all supplied by Applied Biosystems).

Fragment analysis electrophoresis: PCR products are subjected to a post PCR clean-up prior to electrophoresis. Two microliters of cleaned PCR product is mixed with the appropriate internal size

standard (Applied Biosystems) and Hi-Di prior to denaturing for 5 minutes at 95°C. Directly after heating the samples are placed on ice for 5 minutes. Electrophoresis is performed on either an ABI3500xl or an ABI3730xl using a 50cm Capillary array and POP7 (all supplied by Applied Biosystems).

WHOLE EXOME SEQUENCING STANDARD OPERATING PROCEDURES

GENOME DNA QUALITY CONTROL

DNA from these samples were assessed for purity on the NanoDrop™ ND-1000 spectrophotometer (ThermoFisher)

Scientific), using a low Tris-EDTA buffer. The double stranded DNA (dsDNA) fraction in each sample was determined on the Qubit 4.0 fluorometer (ThermoFisher Scientific) using the Qubit™ 1x dsDNA HS assay kit according to the manufacturer's protocol, MAN0017455 REVA.0. Genomic quality scores (GQS) were determined on the LabChip® GXII Touch (PerkinElmer, Waltham, MA, USA), using the DNA Extended Range Chip and Genomic DNA Reagent Kit (PerkinElmer) according to the manufacturer's protocol, CLS140166, Rev. D (Table 1). The GQS is a value between 0 and 5, with zero indicating the lowest quality DNA and five high-quality, intact gDNA.

EXOME LIBRARY PREPARATION

The Ion AmpliSeq™ Exome RDY Kit was used to prepare whole human exome libraries from 100ng input gDNA for sample KB025, KB038 and KB097. Targets were amplified on the SimplyAmp Thermal Cycler using the *Ion AmpliSeq™ Exome RDY Panel* and the *Ion AmpliSeq™ Library Kit Plus* according to the protocol, MAN0010084, REV E.0. Following exome target amplification across 12 primer pools, the products were combined and primer sequences partially digested in preparation for adapter ligation. *IonCode™ Barcode Adapters* were used to generate libraries from project samples with the respective barcodes noted in Table2. Adapter ligated exome libraries were purified *with Agencourt™ AMPure™ XP reagent* and eluted in 50ul low TE buffer.

LIBRARY QUANTIFICATION

Exome library quantification was performed using the *Ion Library TaqMan Quantitation Kit* (MAN0015802 Rev B.0).

Briefly, qPCR reaction volumes were scaled to a final volume of 10ul and amplification was performed using the StepOnePlus™ Real-time PCR system (ThermoFisher Scientific). The molarity of sample libraries is noted in Table2. Table2. Library ID, quantity and run combination.

TEMPLATE PREPARATION AND ENRICHMENT

Libraries were diluted to a target concentration of 20pM. The diluted, barcoded, exome libraries were combined in equimolar amounts for sequencing template preparation using the Ion 550™ Chef Kit. The pooling strategy for the exome libraries are shown in Table2, with the specific run ID as reference. Twenty-five microliters of diluted, pooled library was loaded onto the Ion Chef liquid handler for template preparation and enrichment using Ion 550™ Chef *Reagents*, *Solutions* and *Supplies* according to the protocol, MAN0017275, REVB.0. Enriched ion sphere particles were loaded onto an *Ion 550™ Chip*.

SEQUENCING

Massively parallel sequencing was performed on the Ion S5™ system using Sequencing *Solutions,* and *Reagents* according to the protocol, MAN0017275, REVB.0.

SEQUENCE ANALYSIS

Flow space calibration, base caller analysis, reference genome alignment (CEU), coverage analysis and variant calling was performed using standard analysis parameters in the Torrent Suite Version 5.10.1 Software.

REAGENTS & CONSUMABLES

Item	Manufacturer	Cat No	LOT	Exp. Date
Qubit 1x dsDNA HS assay kit	ThermoFisher Scientific	Q33231	2020112	-
DNA Extended Range LabChip	Perkin Elmer	CLS138948	S993B	2019-05-19
Genomic DNA Reagent Kit	Perkin Elmer	CLS760685	WH07RK01	2019-05-06
Ion AmpliSeq™ Exome RDY Panel, 1x8	ThermoFisher Scientific	4489838	1802026	2021-03-15
Ion AmpliSeq™ Library Kit Plus	ThermoFisher Scientific	4488990	1954919	2020-08-31
Ion Code™ Barcode Adapters Kit	ThermoFisher Scientific	A29748	1702002	2019-03-15
Agencourt™ AMPure™ XP Kit	Beckman Coulter	A63881	170620000	2019-11-15
Ion Library TaqMan™ Quantitation Kit	ThermoFisher Scientific	4468802	00693467	2019-03-31
Ion S5™ 550™ Chef Reagents	ThermoFisher Scientific	A34540	2061837	2019-10-31
Ion S5™ 550™ Chef Solutions	ThermoFisher Scientific	A36410	2061836	2019-09-30
Ion S5™ 550™ Chef Supplies	ThermoFisher Scientific	A27755	2006392	2020-08-31
Ion S5™ 550™Chip	ThermoFisher Scientific	A34537	QRY615E	2019-08-31
Ion S5™ 550™ Sequencing Solutions	ThermoFisher Scientific	100031091	1977159	2019-06-30
Ion S5™ 550™ Sequencing Reagents	ThermoFisher Scientific	INS1012841	1977218	2019-05-31

Standard operating protocol for LR-PCR using PrimeStar GLX

	1X (µL)	Final concentration
DNA (10ng/µL)	4.0	
PrimeStar GLX buffer (5X)	4.0	
dNTPs mix (2.5µM each)	1.6	200µM
Forward primer (10µM)	0.4	0.2µM
Reverse primer (10µM)	0.4	0.2µM
PrimeStar GLX polymerase (1.25U/µL)	0.4	0.5U/µL
Water	9.2	
Total	20.0	

PCR Conditions

98℃ 10 sec

55 or 60℃ 15 sec 30 cycles

68℃ 1 min/kb

} [3-step PCR]

APPENDIX III

CONFERENCE PRESENTATIONS

Okunola AO, Moremi K, Baatjes K, Zemlin A, Erasmus RT, Kotze MJ. Vitamin D levels in estrogen receptor-positive breast cancer patients treated with aromatase inhibitors. International Congress of Clinical Chemistry and Laboratory Medicine. 22-25 October 2017. Durban International Convention Centre, Durban, South Africa.

Okunola AO, Sawe RT, Baatjes KJ, Zemlin AE, Erasmus RT, Kotze MJ. High prevalence of vitamin D deficiency in postmenopausal breast cancer patients treated with aromatase inhibitors at Tygerberg Hospital. CANSA Research in Action Conference, 3-5 July 2018. University of Pretoria, South Africa.

Okunola AO, Sawe RT, Baatjes KJ, Zemlin AE, Erasmus RT, Kotze MJ. Screening of vitamin D-related genes in postmenopausal breast cancer patients using whole exome sequencing. 56th International Congress of the Federation of South African Societies of Pathology (FSASP). PathCape 2018. 16 - 18 August 2018 Spier Stellenbosch, South Africa.

Okunola AO, Sawe RT, Baatjes KJ, Zemlin AE, Erasmus RT, Kotze MJ. Investigation of the role of vitamin D metabolism in South African breast cancer patients using whole exome sequencing. St. Gallen International Breast Cancer Conference, 20-23 March 2019. Vienna, Austria.

Okunola AO, Sawe RT, Baatjes KJ, Zemlin AE, Erasmus RT, Kotze MJ. *CDH1* gene variants identified in breast cancer patients stratified by vitamin D status and tumour histopathology. 18th Biennial Southern African Society for Human Genetics Congress, 3 – 6 August 2019, Century City Conference Centre, Cape Town, South Africa.

Peeters AV, **Okunola AO,** Engelbrecht C, Schoeman M, Kotze MJ. Detection of large deletions in the *BRCA1/2* genes using a combination of whole-exome and MinION-based long-range nanopore sequencing. 18th Biennial Southern African Society for Human Genetics Congress, 3 – 6 August 2019, Century City Conference Centre, Cape Town, South Africa.

Kotze MJ, **Okunola AO,** Peeters AV, van der Merwe N, Torrorey-Sawe R, Baatjes KJ, Conradie M and Schneider JW. Integrating tumour pathology with molecular and biochemical genetics towards individualised treatment of breast cancer. Faculty of Medicine and Health Sciences 63rd Annual Academic Day. 21 August 2019.

WORKSHOP ATTENDED

Next Generation Sequencing Workshop, 13 - 20 April 2018. Wellcome Genome Campus, Hinxton, United Kingdom.