Hereditary Breast Cancer in Selected Populations: Screening *RECQL* Gene in Ontario, Canada, and Exploring *BRCA1/2* and *PALB2* in the Caribbean

by

Humayun Ahmed

A thesis submitted in conformity with the requirements for the degree of Master of Science

Institute of Medical Science

University of Toronto

© Copyright Humayun Ahmed 2018
Hereditary Breast Cancer in Selected Populations: Screening \textit{RECQL} Gene in Ontario, Canada, and Exploring \textit{BRCA1/2} and \textit{PALB2} in the Caribbean

Humayun Ahmed
Master of Science
Institute of Medical Science
University of Toronto
2018

Abstract

Hereditary breast cancer (\textit{i.e.}, cancer due to an inherited mutation) is imperfectly understood: while 5 to 10\% of breast cancer is hereditary, only 50\% of breast cancer susceptibility has been explained. Increased hereditary breast cancer understanding fosters high-risk patient identification, prevention, early diagnosis, and targeted treatment. To achieve increased knowledge of the hereditary breast cancer landscape, both generation of new information and implementation of prior knowledge in new settings are required. Accordingly, this project had two goals: Firstly, we screened \textit{RECQL}, a new susceptibility gene candidate, among Ontario breast cancer patients. Secondly, we examined three established genes – namely, \textit{BRCA1}, \textit{BRCA2}, and \textit{PALB2} – among breast cancer patients in Trinidad and Tobago and Jamaica. Overall, we found that further work was needed to validate \textit{RECQL} in a heterogeneous population, and that Caribbean screening and prevention could be ameliorated by the inclusion of certain susceptibility genes on gene panels.
Acknowledgements

There are numerous individuals and organizations without whom the completion of this thesis would not be possible.

Firstly, I would like to thank Dr. Mohammad Akbari, my direct supervisor, who has supported me daily throughout the last two years. Mohammad, thank you for your patience, your optimism, your belief in my ability to learn new things, your dedicated and meticulous advising on all of our projects, your support in grant searching, and your mentorship. I look forward to working with you in the future, and am deeply grateful for your presence in my professional life.

Secondly, I would like to thank Dr. Steven Narod, who heads the Familial Breast Cancer Research Unit at Women’s College Hospital and Research Institute alongside Dr. Akbari. Dr. Narod consistently challenged me with novel projects and discussions, and has been instrumental to my professional development within research.

I would also like to thank my advisory committee, which included Drs. Steven Gallinger and Jordan Lerner-Ellis, both of whom were very supportive during my advisory committee presentations, provided very constructive, clear feedback following these presentations, and made themselves consistently available to me for consultations despite their many commitments.

This work would also, of course, not have been possible without the generous help of the entirety of the Familial Breast Cancer Research Unit at the Women’s College Hospital and Research Institute. Special thanks is due to Robert Royer, who provided me with a solid, meticulous introduction to the foundations of the protocols employed in breast cancer genetics when I first joined the Familial Breast Cancer Research Unit; Shiyu Zhang, who contributed an abundance of wet lab help to this project; Victoria Sopik, who was
always available for comments and advice on joint papers; Erin Sellars, who carried out the Sanger sequencing necessary to validating this project’s findings, coordinates our wet lab operations, and was an invaluable resource where learning new protocols were concerned; Ellen MacDougall, who consistently aligns all of our colleagues’ schedules; and Marcia Llacuachaqui, who was always willing to lend an encouraging word and a helping hand.

The Institute of Medical Science’s Master of Science (Medical Science) program is another institution that deserves my sincerest gratitude: all of the IMS faculty and staff with whom I have had the pleasure of interacting through my courses have solely been helpful to the definition of my career and the execution of this publication. Special thanks is due to Drs. Lucy Osborne and Steven Scherer, who taught MSC2010, a course which immersed me in a range of key genetics and epidemiology concepts, and to Hazel Pollard, who was always available to assist me with logistics.

Without external financial support, this work also would not have been possible. I would thus like to extend my deepest thanks to the Province of Ontario and the University of Toronto in awarding me the Ontario Graduate Scholarship (OGS), the Ministry of Training, Colleges, and Universities and the University of Toronto for awarding me the Queen Elizabeth II Graduate Scholarship in Science & Technology (QEII-GSST) and the Institute of Medical Science for awarding me the Institute of Medical Science Entrance Award.

Personally speaking, I would, of course, also like to thank my family and friends, without whom it would not have been possible for me to complete this project.
Finally, I would like to thank all those women who participated in this project for sharing their lives and their battles with us, in the hopes of furthering both their own treatments and the future medical treatment possible for other women.
Table of Contents

ACKNOWLEDGEMENTS ........................................................................................................III

LIST OF ABBREVIATIONS .....................................................................................................VIII

INTRODUCTION AND OBJECTIVES ...................................................................................... 1
1.1 INTRODUCTION .................................................................................................................. 1
1.2 OBJECTIVES ....................................................................................................................... 5
  1.2.1 Project One: Screening RECL in Ontario, Canada ......................................................... 5
  1.2.2 Project Two: BRCA1, BRCA2, and PALB2 in Trinidad and Tobago ......................... 6
  1.2.3 Project Three: BRCA1, BRCA2, and PALB2 in Jamaica ............................................. 6

INTRODUCTION TO BREAST CANCER ............................................................................. 7
  2.1 Types of Breast Cancer: Sporadic, Hereditary, and Familial ............................................. 7
  2.2 RISK FACTORS ................................................................................................................ 8
    2.2.1 Pregnancy and Childbirth ........................................................................................... 8
    2.2.2 Age ............................................................................................................................... 8
    2.2.3 Behavioral and Environmental Influences ................................................................... 9
  2.3 GENETIC CONSIDERATIONS ......................................................................................... 11
    2.3.1 Molecular and Genetic Subtypes ............................................................................... 11
  2.4 SCREENING AND DIAGNOSTICS ............................................................................... 13
    2.4.1 Screening .................................................................................................................. 13
    2.4.2 Diagnosis .................................................................................................................. 15
  2.5 MANAGEMENT ................................................................................................................. 17
    2.6 Treatment Limitations and Future Directions ............................................................... 18

HEREDITARY BREAST CANCER .................................................................................. 21
  3.1 HEREDITY AND BREAST CANCER .............................................................................. 21
  3.2 UNDERSTANDING HEREDITARY BREAST CANCER: KEY TERMINOLOGY ............. 25
    3.2.1 Mutation- and Gene-Level Terms ........................................................................... 26
    3.2.2 Tumor-Level Terms ................................................................................................. 32
  3.3 BREAST CANCER SUSCEPTIBILITY GENES: GENERAL FUNCTIONS AND INVOLVED
    PATHWAYS ....................................................................................................................... 33
    3.3.1 Cell Division ............................................................................................................. 33
    3.3.2 DNA Damage Recognition and Response ................................................................. 34
    3.3.3 DNA Repair ............................................................................................................... 34
    3.3.4 Other Functions: Cell Survival Post-Damage, Transcription, Apoptosis,
        Senescence, Autophagy, and Angiogenesis .................................................................... 38
    3.3.5 Joint Functions: Gene Interactions ........................................................................... 38
    3.3.6 Two Cancerous Categories: Oncogenes and Tumor Suppressors ................................ 39
  3.4 BREAST CANCER SUSCEPTIBILITY GENES: KNOWN MUTATION TYPES ............... 40
    3.4.1 Structural Types ........................................................................................................ 40
    3.4.2 Functional Types ...................................................................................................... 44
  3.5 KNOWN BREAST CANCER SUSCEPTIBILITY GENES .............................................. 45
    3.3.1 Breast cancer 1 (BRCA1) and breast cancer 2 (BRCA2) ........................................... 45
3.3.2 Tumor protein 53 (TP53, p53)................................................................. 54
3.3.3 Phosphatase and tensin homolog (PTEN) ............................................... 57
3.3.4 Other high-penetrance genes .................................................................. 58
3.3.5 Moderate-penetrance genes ................................................................... 59
3.3.6 Plausible and former candidate genes ....................................................... 61
3.3.7 Low-penetrant loci .................................................................................. 63

SCREENING RECQL GENE AMONG BREAST CANCER PATIENTS IN
ONTARIO, CANADA .......................................................................................... 64
4.1 RECQL: INTRODUCTION ........................................................................... 64
  4.2.1 Introduction .......................................................................................... 64
  4.2.2 Structure ............................................................................................. 64
  4.2.3 Function ................................................................................................ 65
4.3 RECQL: BREAST CANCER SUSCEPTIBILITY GENE CANDIDACY IN SELECTED
POPULATIONS ................................................................................................. 67
  4.3.1 Our Lab’s Prior Paper .......................................................................... 67
  4.3.2 Other Papers ........................................................................................ 70
4.4 OUR PROJECT: SCREENING RECQL IN ONTARIO POPULATION .............. 72
  4.4.1 Rationale ............................................................................................. 72
  4.4.2 Objective .............................................................................................. 74
  4.4.3 Methodology ......................................................................................... 74
  4.4.4 Results ................................................................................................. 80
  4.4.5 Discussion ............................................................................................ 86

OLD GENES, NEW CONTEXT: BRCA1, BRCA2, AND PALB2 IN THE
CARIBBEAN ....................................................................................................... 90
5.1 CANCER IN SELECTED POPULATIONS .................................................... 90
5.2 BREAST CANCER IN THE CARIBBEAN .................................................. 91
5.3 PROJECT ONE: A SURVEY OF BRCA1, BRCA2, AND PALB2 MUTATIONS IN
WOMEN WITH BREAST CANCER IN TRINIDAD AND TOBAGO ................. 92
  5.3.1 Rationale ............................................................................................. 92
  5.3.2 Objectives ............................................................................................ 93
  5.3.3 Methods ................................................................................................ 93
  5.3.4 Results ................................................................................................. 95
  5.3.5 Discussion ............................................................................................ 102
5.4 PROJECT TWO: A HIGH FREQUENCY OF PALB2 MUTATIONS IN JAMAICAN
PATIENTS WITH BREAST CANCER ................................................................. 103
  5.4.1 Rationale ............................................................................................. 103
  5.4.2 Objectives ............................................................................................ 104
  5.4.3 Methods ................................................................................................ 104
  5.4.4 Results ................................................................................................. 105
  5.4.5 Discussion ............................................................................................ 109

FUTURE DIRECTIONS ...................................................................................... 111
6.1 RECQL ........................................................................................................ 111
6.2 BRCA1, BRCA2, & PALB2 IN THE CARIBBEAN ....................................... 113
REFERENCES ................................................................................................ 114
List of Abbreviations

ABRAXAS  BRCA1-A Complex Subunit
ACACA    Acetyl-CoA Carboxylase Alpha
ATM      Ataxia Telangiectasia Mutated
ATR      Ataxia Telangiectasia and Rad3-Related Protein
AURKA    Aurora Kinase A
BAM      Binary Alignment Map
BAP1     BRCA1-Associated Protein 1
BARD 1   BRCA1-Associated RING Domain 1
BASC     BRCA1-associated genome surveillance complex
BER      Base excision repair
BIR      Break-induced repair
BLM      Bloom syndrome RecQ-Like Helicase
BMI      Body mass index
BRCA 1   Breast Cancer 1
BRCA 2   Breast Cancer 2
BRCC     BRCA1/BRCA2-containing complex
BRCC3    BRCA1/BRCA2-Containing Complex Subunit 3
BRCC45   BRCA1/BRCA2-Containing Complex Subunit 45
BRCT     Breast cancer carboxy-terminal domain
BRIPI     BRCA1-Interacting Protein C-Terminal Helicase 1
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSE</td>
<td>Breast Self-Examination</td>
</tr>
<tr>
<td>BWA</td>
<td>Burrows Wheeler Aligner</td>
</tr>
<tr>
<td>CARPHA</td>
<td>The Caribbean Public Health Agency</td>
</tr>
<tr>
<td>CBE</td>
<td>Clinical Breast Examination</td>
</tr>
<tr>
<td>CCDC98</td>
<td>Coiled-Coil Domain Containing 98</td>
</tr>
<tr>
<td>CCND1</td>
<td>Cyclin D1</td>
</tr>
<tr>
<td>CDC</td>
<td>Center of Disease Control</td>
</tr>
<tr>
<td>CDH1</td>
<td>Cadherin 1</td>
</tr>
<tr>
<td>CHEK2</td>
<td>Checkpoint Kinase 2</td>
</tr>
<tr>
<td>CNA</td>
<td>Copy number alterations</td>
</tr>
<tr>
<td>CNV</td>
<td>Copy number variations</td>
</tr>
<tr>
<td>COBRA1</td>
<td>Negative Elongation Factor Complex Member B (NELFB)</td>
</tr>
<tr>
<td>CT</td>
<td>Computed Tomography</td>
</tr>
<tr>
<td>CTC</td>
<td>Circulating Tumor Cells</td>
</tr>
<tr>
<td>dbSNP</td>
<td>Single Nucleotide Polymorphism Database</td>
</tr>
<tr>
<td>DCIS</td>
<td>Ductal carcinoma in-situ</td>
</tr>
<tr>
<td>DSB</td>
<td>Double strand break</td>
</tr>
<tr>
<td>DSBR</td>
<td>Double strand break repair</td>
</tr>
<tr>
<td>ER</td>
<td>Estrogen receptor</td>
</tr>
<tr>
<td>ERBB2</td>
<td>Human Epidermal Growth Factor Receptor 2</td>
</tr>
<tr>
<td>ESR1</td>
<td>Estrogen Receptor 1</td>
</tr>
<tr>
<td>ETV6</td>
<td>ETS Variant 6</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>ExAC</td>
<td>The Exome Aggregation Consortium</td>
</tr>
<tr>
<td>EXD2</td>
<td>Exonuclease 3'-5' Domain Containing 2</td>
</tr>
<tr>
<td>FANCD1</td>
<td>Breast Cancer 2</td>
</tr>
<tr>
<td>FANCN</td>
<td>Fanconi Anemia, Complementation Group N</td>
</tr>
<tr>
<td>FGFR2</td>
<td>Fibroblast Growth Factor Receptor 2</td>
</tr>
<tr>
<td>GATK</td>
<td>Genome Analysis Toolkit</td>
</tr>
<tr>
<td>GWAS</td>
<td>Genome-wide association study</td>
</tr>
<tr>
<td>H19</td>
<td>Non-Protein Coding RNA 8</td>
</tr>
<tr>
<td>HBOC</td>
<td>Hereditary breast and ovarian cancer</td>
</tr>
<tr>
<td>HDAC2</td>
<td>Histone Deacetylase 2</td>
</tr>
<tr>
<td>HDM2</td>
<td>Human MDM2 Proto-Oncogene</td>
</tr>
<tr>
<td>HER 2</td>
<td>Human Epidermal Growth Factor Receptor 2</td>
</tr>
<tr>
<td>HOXA5</td>
<td>Homeobox A5</td>
</tr>
<tr>
<td>HRR</td>
<td>Homologous recombination repair</td>
</tr>
<tr>
<td>IDC</td>
<td>Invasive ductal carcinoma</td>
</tr>
<tr>
<td>IHC</td>
<td>Immunohistochemistry</td>
</tr>
<tr>
<td>ILC</td>
<td>Invasive lobular carcinoma</td>
</tr>
<tr>
<td>IR</td>
<td>Ionizing radiation/infrared radiation</td>
</tr>
<tr>
<td>IRB</td>
<td>Institutional review board</td>
</tr>
<tr>
<td>ISH</td>
<td>In-situ hybridization</td>
</tr>
<tr>
<td>ITPR2</td>
<td>Inositol 1,4,5-Trisphosphate Receptor Type 2</td>
</tr>
<tr>
<td>KPNA4</td>
<td>Karyopherin Subunit Alpha 4</td>
</tr>
</tbody>
</table>
LCIS  Lobular carcinoma in situ

*LISP1*  Liver-Specific Protein 1

LOH  Loss of heterozygosity

*LSP1*  Lymphocyte-Specific Protein 1

MAF  Minor allele frequency

*MAP3K*  Mitogen-Activated Protein Kinase Kinase Kinase

*MAP3K1*  Mitogen-Activated Protein Kinase Kinase Kinase 1

*MLH1*  MutL Homolog 1

MMEJ  Microhomology-mediated end-joining

MMR  Mismatch repair

MRI  Magnetic resonance imaging

*MRN*  MRE11-RAD50-NBN complex

*MSH*  MutS Homolog

*MSH2*  MutS Homolog 2

*MSH6*  MutS Homolog 6

mTOR  Mechanistic target of rapamycin

*MYC*  Avian Myelocytomatosis Viral Oncogene Homolog

*NBA1*  BRISC and BRCA1 A Complex Member 1

NBN  Nibrin

NER  Nucleotide excision repair

NGS  Next-generation sequencing

NHEJ  Non-homologous end-joining repair
NHLBI  National Heart, Lung, and Blood Institute

NRG1  Neuregulin 1

OR  Odds ratio

PALB2  Partner and Localizer of BRCA2

PARP  Poly(ADP-Ribose) Polymerase

PCLAF  PCNA clamp-associated factor

PET  Positron-emission tomography

PI3K  Phosphatidylinositol-4,5-bisphosphate 3-kinase

PIK3CA  Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit

PMS2  PMS Homolog 2

PR  Progesterone receptor

PTEN  Phosphatase and Tensin Homolog

RAD50  RAD50 Double Strand Break Repair Protein

RAD51  RAD51 Recombinase

RAD51C  RAD51 Paralog C

RAD51D  RAD51 Paralog D

RAD51LI  RAD51 Paralog B

RAP80  Ubiquitin Interaction Motif Containing 1

RASSF1A  Ras Association Domain Family Member 1

RECQL  RecQ Helicase-Like

RFC  Replication Factor C

SAM  Sequence Alignment Map
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCD</td>
<td>Serine cluster domain</td>
</tr>
<tr>
<td>SDSA</td>
<td>Synthesis-dependent strand annealing</td>
</tr>
<tr>
<td>SIPA1</td>
<td>Signal Induced Proliferation Associated 1</td>
</tr>
<tr>
<td>SMC1A</td>
<td>Structural Maintenance of chromosome 1</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
<tr>
<td>SSB</td>
<td>Single strand break</td>
</tr>
<tr>
<td>STK 11</td>
<td>Serine/Threonine Kinase 11</td>
</tr>
<tr>
<td>TBM</td>
<td>Tandem base mutation</td>
</tr>
<tr>
<td>TCGA</td>
<td>The Cancer Genomic Atlas</td>
</tr>
<tr>
<td>TGFB1</td>
<td>Transforming Growth Factor Beta 1</td>
</tr>
<tr>
<td>TMS1</td>
<td>Target of Methylation-Induced Silencing 1</td>
</tr>
<tr>
<td>TNBC</td>
<td>Triple-negative breast cancer</td>
</tr>
<tr>
<td>TNM</td>
<td>Tumor, Node, Metastasis</td>
</tr>
<tr>
<td>TNRC9</td>
<td>Trinucleotide Repeat-Containing Gene 9 Protein</td>
</tr>
<tr>
<td>TOX3</td>
<td>TOX High Mobility Group Box Family Member 3</td>
</tr>
<tr>
<td>TP53</td>
<td>Tumor Protein 53</td>
</tr>
<tr>
<td>UBXN1</td>
<td>UBX Domain Protein 1</td>
</tr>
<tr>
<td>UST</td>
<td>Uronyl 2-Sulfotransferase</td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascular Endothelial Growth Factor</td>
</tr>
<tr>
<td>WCRI</td>
<td>Women’s College Research Institute</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>WRN</td>
<td>Werner Syndrome RecQ-Like Helicase</td>
</tr>
</tbody>
</table>
*XRCC1*  X-ray Repair Cross-Complementing Protein 1

*ZMIZ1*  Zinc Finger, MIZ-Type Containing 1
Chapter 1

Introduction and Objectives

1.1. Introduction

Breast cancer is the most common cancer in women around the world\(^1\), and the second leading cause of cancer-related death in Canadian women\(^2\). In 2016 alone, an estimated 25,700 Canadian women were diagnosed with breast cancer\(^3\), and an estimated 4,900 Canadian women died from breast cancer\(^4\). Worldwide, the number of affected women was likely far higher: according to the most recent data available, which is from 2012, more than 1.7 million new breast cancer cases and 500,000 breast cancer deaths occur among women worldwide every year, with developed countries generally having higher rates than developing countries\(^5\).

Approximately 5 to 10% of all breast cancer cases are hereditary cancer cases\(^6\), meaning that the individual in question inherited the causative breast cancer mutation from his or her parents. Another 10% of all breast cancer cases are familial cancer cases\(^7\), meaning that though their causal genetic mutations are not known, the patients have a history of breast cancer in their immediate relatives. Both inheriting a breast cancer-relevant mutation and having affected family members are significant breast cancer risk factors: mutations in certain high-penetrance breast cancer susceptibility genes may increase a woman’s lifetime breast cancer risk by as much as 80%\(^8\), while the presence of affected individuals in a woman’s family can make her lifetime breast cancer risk up to twice as high as that of an individual without affected relatives, even independently of any inherited mutations\(^9\). Identifying breast cancer susceptibility genes
and individuals who are at risk of developing breast cancer due to a mutation in those genes is important from several perspectives. Firstly, knowing an individual’s genetic make-up and isolating mutations in known susceptibility genes helps us take preventive measures to reduce cancer risk: for example, mutations in \textit{BRCA1/2} are associated with an up to 80% life-time risk of developing cancer; patients with such mutations can be advised to seek prophylactic measures such as double mastectomy, which has been shown breast cancer risk by up to 95\%\textsuperscript{10}. Secondly, knowing about breast cancer susceptibility genes allows us to offer intensive screening to high-risk individuals (\textit{e.g.}, those with strong family history), allowing us to detect breast cancer at its early stages, which ameliorates the chances of treatment succeeding. Finally, many biomarkers are predictive not solely of breast cancer development, but also of the response to treatment – and consequently also the prognosis – that a mutation carrier may have. As an example, individuals with \textit{BRCA1} mutations are more likely than \textit{BRCA2} mutation carriers to be both basal-like or triple negative (TN), which is also typically associated with a worse prognosis\textsuperscript{11}. If we are aware of the associations between certain biomarkers and certain prognoses, then we can better predict not solely which patients will get breast cancer, but also which patients have more aggressive cancers requiring more intensive treatments. Also, genes – if they are found to be causal – may potentially be targeted via novel drugs developed specifically to cater to them. For example, \textit{BRCA1} mutation carriers appear to be more sensitive to cisplatinum therapy\textsuperscript{12}.

It is clear, then, that susceptibility gene discovery is integral to oncology for a range of reasons. Given the importance of discovery, and the relatively recent advent of next-generation sequencing (NGS), a form of high-throughput modern sequencing in
which massive amounts of genetic information can be processed in parallel, one would expect that gene discovery would be proceeding more quickly than ever before, considering that more genetic material can now be analyzed more cost- and time-efficiently. However, the vast majority of key breast cancer susceptibility genes (e.g., TP53, BRCA1, ATM, BRCA2, CDH1, CHEK2, PTEN, BARD1, NBN, and STK11) were discovered by the 1990s\textsuperscript{13,14,15}; by contrast, since 2000, only one significant moderate-penetrance breast cancer susceptibility gene – namely, \textit{PALB2} – has been discovered\textsuperscript{16}. This suggests that the vast majority of high-penetrance genes with relatively high mutation frequencies in the general population have likely already been discovered. Next-generation sequencing, which is gaining popularity, permits rapid, in-depth surveillance of the entirety of the genome, but few high-penetrance genes have been found. Any undiscovered genes would likely either be low-penetrance genes whose effects are only expressed in a small percentage of mutation carriers, or exceedingly rare variants of high-penetrance genes.

In search of just such rare high-penetrance breast cancer susceptibility genes, our lab conducted whole-exome next-generation sequencing of two populations (from Poland and Quebec) and reported a novel breast cancer susceptibility gene called \textit{RECQL}\textsuperscript{17}. While many genes were found to have rare, truncating variants in either the Polish or the Quebecois population, only had multiple rare truncating mutations in both populations. It should also be noted that, in addition to being the only gene that harbored multiple rare truncating mutations across both populations (and, consequently, the only gene that yielded a promising odds ratio), \textit{RECQL} is functionally likely to be a susceptibility gene. Like the other four RecQ helicases, \textit{RECQL} likely plays a role in one or more important
cellular functions related to DNA repair, replication, recombination, and transcription. It is also a potential regulator of genomic integrity: when it is knocked out in mice embryonic fibroblasts, a range of genomic complications occur, including spontaneous chromosomal breakage, aneuploidy, and translocations; when both mouse and human cells are depleted of it, they grow less, become more sensitive to ionizing radiation (IR), experience a high rate of spontaneous sister chromatid exchange, and develop more double-strand DNA breaks (dsDNA breaks).\textsuperscript{18,19,20} \textit{RECQL} has also been shown to be involved in non-homologous end joining DNA repair (NHEJ), a pathway in which double-strand DNA breaks are fixed via the broken ends’ being directly ligated to one another without the need for a homologous template, and with lengthening telomeres independently of telomerase, the enzyme usually responsible for telomere elongation.\textsuperscript{21} For all these reasons, \textit{RECQL} is thought to be a tumor suppressor: a gene that typically discourages tumor formation, notably by preventing replication forks from stalling and breaking, and whose loss of function via mutations thus leads to increased cancer susceptibility.\textsuperscript{22}

On these bases, our team identified \textit{RECQL} as a breast cancer susceptibility gene candidate. However, this initial discovery-oriented study was conducted in two populations that are not indicative of the general population of Canada. Both Poland and Quebec are regions with founder populations: populations that share genetic characteristics because of shared ancestry. As such, even a very high coincidence of a rare truncating mutation with breast cancer risk in such a population may not be probative of that mutation’s predictive value in the general population, as this causative mutation may well be a founder mutation: a mutation primarily restricted to the smaller, more
homogeneous populations in question. Follow-up work concerning *RECQL* in the context of the general Canadian population was needed; this is where the first project that was carried out here becomes relevant.

At the same time, as much as discovering and validating new breast cancer susceptibility genes is of relevance, it is also important to examine existing breast cancer susceptibility genes in populations with high breast cancer incidence, and about which we have little established hereditary breast cancer information. The Caribbean, for example, is a region with high breast cancer incidence, but relatively little in the way of formally instituted national screening policies for unselected patients, many of whom could potentially benefit from genetic testing and related prevention and treatment strategies. To this end, we also carried out two projects geared at examining the frequency of *BRCA1*, *BRCA2*, and *PALB2* mutations in Trinidad and Tobago and Jamaica, two Caribbean countries with a sparse genetic testing regulation landscape, but very high incidences of breast cancer.

### 1.2. Objectives

**1.2.1 Project One: Screening *RECQL* in Ontario, Canada**

Objective 1: Screen the entire coding exons of the *RECQL* gene, plus 20 base pairs of the intronic regions at each end of the exons, among 2,859 breast cancer cases and 929 healthy women from Ontario to locate *RECQL* mutation carriers. Emphasis will be placed on locating variants that are potentially pathogenic (i.e. truncating variants).

Objective 2: Estimate the odds ratio of developing breast cancer and possibly the life-time risk of developing breast cancer associated with carrying a *RECQL* mutation.
1.2.2 Project Two: BRCA1, BRCA2, and PALB2 in Trinidad and Tobago
Objective: Examine the frequency of BRCA1/2 and PALB2 mutations in
unselected Trinidad and Tobago breast cancer cases to determine the contribution of
hereditary cases to that population.

1.2.3 Project Three: BRCA1, BRCA2, and PALB2 in Jamaica
Objective: Examine the frequency of BRCA1/2 and PALB2 mutations in
unselected Jamaican breast cancer cases to determine the contribution of hereditary cases
to that population.
Chapter 2

Introduction to Breast Cancer

2.1 Types of Breast Cancer: Sporadic, Hereditary, and Familial

Breast cancer can be caused by both germline and somatic mutations. Germline mutations are mutations that are mostly inherited from one’s parents, and are thus present in the genome since birth\(^23\). By contrast, somatic mutations are mutations that have not been inherited from one’s parents, but have instead been acquired by the cell sporadically through DNA replication and via other influences over the course of a lifetime\(^24\). If breast cancer is caused by germline mutations, it is referred to as hereditary or inherited breast cancer\(^25\). If it is caused by somatic mutations, it is referred to as sporadic breast cancer.

Further, if a given family has a history of breast cancer, then familial breast cancer is said to be present. A family history is considered strong if a person’s family includes two affected first-degree relatives (\textit{i.e.}, parents, children, or siblings); a strong family history is associated with a higher risk of breast cancer, though not necessarily with hereditary breast cancer (\textit{i.e.}, it is possible to have familial but not inherited breast cancer)\(^26\). If there is any first-degree family history, then there is usually a two-fold increased risk of breast cancer\(^27\); the higher the number of affected relatives, the higher the relative risk\(^28\); finally, the less close the relationship between the relatives, the lower the risk (\textit{e.g.}, risk is only 1.5-fold for second-degree relatives compared to two-fold for first-degree relatives)\(^29\). When an individual has a family history of breast cancer but no known genetic anomalies associated with the cancer, the cancer is said to be familial. When, on the other hand, an individual has a family history of breast cancer and known mutations in breast cancer susceptibility genes, the cancer is said to be hereditary.
2.2 Risk Factors

2.2.1 Pregnancy and Childbirth

Pregnancy and childbirth affect a woman’s risk of developing breast cancer. Parity, or the number of pregnancies a woman has had, modifies breast cancer risk. Nulliparity (i.e., having no pregnancies) is a breast cancer risk factor. Generally, the more pregnancies a woman has had, the less likely she is to develop breast cancer, particularly if the first child is had before 20 years of age (in which case a 50% reduction in life-time breast cancer risk compared to nulliparous women has been observed). However, it may be the case that increased parity, though it decreases the risk of luminal A breast cancer, increases the risk of triple negative breast cancer. This association, however, as well as associations between parity and luminal B and HER2-overexpressing breast cancers, are not presently clear. Additionally, the longer a woman breastfeeds, the lower her risk of developing breast cancer of the basal-like subtype, and the lower her risk of recurrence and breast cancer death. This has important epidemiological implications, given that black women in the U.S. breastfeed significantly less than white women, and experience nearly double the rate of triple-negative breast cancer. Additionally, as has already been mentioned, the age at which a woman first gives birth also affects the association between parity and luminal A breast cancer: specifically, the younger a woman is the first time she gives birth, the more decreased her risk of breast cancer will be, assuming equivalent parity.

2.2.2 Age

Age in general is decidedly a risk factor for breast cancer. The older a woman gets, the more likely she is to get breast cancer, with risk doubling approximately every ten years until menopause, at which point risk continues to increase, but does so
comparatively slowly. However, some breast cancers tend to occur somewhat earlier than certain other cancers would: there is higher incidence of breast cancer than lung cancer overall at younger ages, for example. Breast cancers are frequently categorized on the basis of whether they occur early (“early-onset”, “pre-menopausal”) or late in a person’s life (“late-onset”, “post-menopausal”), and certain types of breast cancers are more likely to occur early than other types. Invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC), for example, tend to affect mostly older women (ages 55 or older), though certain subtypes of IDC (e.g., medullary) as well as LCIS involve earlier onset, while some subtypes (e.g., mucinous) tend to involve unusually late onset.

More specifically, age at menarche and age at menopause are relevant age-related risk factors for breast cancer. Typically, older age at menarche tends to be correlated with a slightly decreased risk of certain more problematic cancers (i.e., triple negative and luminal A, as well as possibly luminal B). However, relatively little has been confirmed about the relationship between HER2-overexpressing cancer and age at menarche. By contrast, older age at menopause appears to be correlated with an increased risk of breast cancer.

2.2.3 Behavioral and Environmental Influences
Certain attributes of a woman’s life history are also risk factors for breast cancer. Body-mass index (BMI), the ratio of a person’s weight to their height, tends to be correlated with decreased risk of luminal A breast cancer and, in some studies, increased risk of triple negative breast cancer when elevated in pre-menopausal women. In post-menopausal women, increased BMI has been associated with an increased risk of breast cancer. In pre-menopausal women, use of oral contraceptives can also influence a woman’s breast cancer risk: specifically, use of oral contraceptives
tends to be correlated with decreased risk of luminal A breast cancer, and increased risk of triple negative breast cancer\textsuperscript{50}. Correspondingly, in post-menopausal women, use of menopausal hormone therapy is associated with an increased risk of breast cancer\textsuperscript{51}.

Finally, where substances are concerned, increased alcohol use (\textit{i.e.,} more than seven drinks per week compared to none) appears to be correlated with a somewhat increased risk of breast cancer\textsuperscript{52}.

Personal factors aside, geography is also a risk factor for breast cancer. Standardized mortality for breast cancer tends to be higher in developed rather than developing countries\textsuperscript{53}, with mortality being highest in Belgium at an age-standardized rate of 111.9 per 100,000\textsuperscript{54}. However, lack of awareness is a serious issue in underdeveloped countries, where it was recently found that both prognostication and survival rates are low due to lack of early diagnosis\textsuperscript{55}. Studies have also suggested that this increased mortality is more due to the environmental and lifestyle factors of a given locality than to the genetics of its natives, considering that breast cancer risk tends to assimilate to the risk of the population that a person has immigrated into within one or two generations\textsuperscript{56}.

Finally, family history of breast cancer is a risk factor, as could be expected given the highly hereditary nature of certain types of breast cancer. If one or more first-degree relatives (\textit{i.e.,} parents, siblings, or children) has had breast cancer, the risk of developing any kind of breast cancer substantially increases: in fact, the risk of breast cancer in individuals with a family history can be up to two times higher than the risk of breast cancer in individuals without a family history.
It is worth mentioning at this point that several associations have been noted between the aforementioned risk factors and luminal A cancer, but that significantly less is known about the associations between these risk factors and other cancer types (*i.e.*, luminal B, HER2-overexpressing, and triple negative).

### 2.3 Genetic Considerations

#### 2.3.1 Molecular and Genetic Subtypes

While breast cancer can be described as per the above phenotypic attributes, breast cancer can – and should – also be classified in terms of its genetics. Segmenting breast cancer into molecular subtypes permits the assignment of treatment options that align well with a specific cancer’s genetic characteristics. For several decades, we have recognized that breast cancer is a highly heterogeneous condition, with variation occurring both between tumors (inter-tumor variation) and within individual tumors (intra-tumor variation)\(^57\). Despite this variation, advances in genetic technologies have made it possible to segment breast cancer into five core molecular subtypes: luminal A, luminal B, luminal B-like, HER2-positive, and basal-like. Cancers are classified into subtypes on the basis of gene expression profiling of the tumor cells\(^58\).

Luminal A breast cancers tend to involve several characteristics. Firstly, luminal A cancers tend to be estrogen receptor-positive (ER+)\(^59\), which means that they typically respond to estrogen, and can thus be managed via drugs that interfere with estrogen production or binding, as is further explained in Chapter Two. In a similar vein, they tend to be progesterone-receptor positive (PR+)\(^60\), which again means they can be affected by therapies that interfere with progesterone production and function. Luminal A cancers also tend to be HER2-negative (HER2-)\(^61\), which means that they do not have amplified
or overexpressed HER2. Finally, luminal A tumors tend to have low quantities of the protein Ki-67 (also known as MKI67), which is a protein that is associated with cell proliferation. All of this may somewhat help us understand why luminal A breast cancers are usually not as prognostically problematic: ER- and PR-positivity implies that these cancers can be treated with the comparatively innocuous hormonal therapy rather than the more damaging chemotherapy, HER2-negativity confers better prognosis, given that HER2-positivity is associated with p53 abnormalities, lymphoid infiltration, and a high mitotic index, and lower quantities of Ki-67 potentially indicate that less cancerous cell proliferation has occurred, explaining the luminal A association with smaller tumor sizes.

Luminal B breast cancers differ from luminal A cancers in a number of ways. Like luminal A cancers, they tend to be both ER-positive and HER2-negative. However, unlike luminal A cancers, they tend to be PR-negative, making them less susceptible to hormonal therapy. They may also have a higher Ki-67 index, which has been found to be correlated with young age at diagnosis, larger tumors, invasion of the lymph nodes, and p53 overexpression (in addition to HER2-positivity and ER-negativity, which luminal B cancers nevertheless lack), and is thus both prognostically disadvantageous and prone to conferring earlier recurrence.

Luminal B-like cancers resemble luminal B breast cancers in that they are ER-positive. However, unlike luminal B cancers, luminal B-like cancers involve overexpression or amplification of HER2, which connotes a poorer prognosis. They also involve any levels of Ki-67 and PR.

HER2-positive cancers, as their name suggests, involve solely HER2-positivity, which involves a poorer prognosis. In some breast cancer types, HER2-positivity also
tends to inversely correlate with ER and PR positivity\textsuperscript{67}, which means that HER2-positive cancers will often need to be treated with chemotherapy rather than with hormonal therapy.

Finally, basal-like cancers are negative where all hormone receptors are concerned: they lack ER, PR, and HER2 receptors (ER-negative, PR-negative, HER2-negative)\textsuperscript{68}. While HER2-negativity would usually connote a positive prognosis, the fact that these cancers are hormone receptor-negative makes them unusually difficult to treat.

These subtypes are important because they have prognostic relevance. Some subtypes respond better to chemotherapy and hormonal therapy than others: intuitively, for example, subtypes lacking certain hormone receptors will not be responsive to hormonal therapy.

2.4 Screening and Diagnostics

2.4.1 Screening

Breast cancer screening involves monitoring healthy women for signs of breast cancer, with the aim of detecting any existing cancers as early as possible. Screening can consist of a range of practices, including awareness-raising, physical examinations, patient history collection, MRI, mammography, and risk assessment\textsuperscript{69}. Breast self-examination (BSE), in which the patient self-examines their breasts and reports changes to their physician, is also a part of the screening process.

Typically, formal screening should be conducted once every one to three years, depending on the age of the individual in question, and involves patient history collection and a clinical breast examination (CBE), which is based on palpation of the breast for symptoms of potential cancer, including lumps, nipple discharge, skin changes, and
asymmetric thickening\textsuperscript{70}. Following a CBE in which potential symptoms are found, individuals’ cancer risks are stratified: a woman may be deemed to be at normal or increased risk of breast cancer. Risk is a function of a range of variables, including family history and mutations, and can be calculated using a range of models (\textit{e.g.}, the empirical Gail model, genetic models)\textsuperscript{71}. A family history of breast cancer in first-degree relatives, for example, confers a two-fold increased risk in breast cancer, and is considered strong if two or more first-degree relatives are affected.

Further, certain mutated genes, as will be discussed in Chapter Four, confer an increased risk of breast cancer; given that a considerable percentage of variation in breast cancer risk is genetic, these need to be factored into risk as well. For example, \textit{TP53}, when mutated, confers a 50 to 60\% breast cancer risk by the age of 45\textsuperscript{72}; \textit{PTEN}, slightly less severely, confers a 30 to 50\% risk by the age of 70\textsuperscript{73}. Generally, women with first-degree relatives below the age of 40 or with two affected first-degree relatives at any age are considered to be a moderate risk\textsuperscript{74}; women with two or more affected first-degree relatives with at least one diagnosed under the age of 50 are considered high-risk\textsuperscript{75}. Finally, note that there are cases in which mutation frequency is dependent on family history: if, for example, a strong family history involving six or more cases is present, then it is probable that at least some of that familial cancer is associated with \textit{BRCA} mutations\textsuperscript{76}. It is also noteworthy that risk of ovarian cancer is 30\% higher in mothers and sisters of individuals with breast cancer\textsuperscript{77,78}, given that the \textit{BRCA} genes are responsible for both ovarian and breast cancer contributions. Risk prediction also involves analyzing some of the lifestyle factors described earlier.
At normal risk, women are encouraged to get a CBE everyone to three years if they are between 20 to 39, and to get an annual CBE as well as a screening mammography if they are 40 years or older. At increased risk, women are told to get annual mammograms and a CBE every six to twelve months if they are over 25 and have previously undergone radiation; if they are under 25, then an annual CBE is recommended. For women aged 35 years and older, a risk assessment tool is available; if these women end up having a five-year cancer risk of 1.7 % or greater, CBEs are recommended every six to 12 months, alongside annual mammography. In certain special cases where risk is unusually high (greater than 20 % lifetime risk, based largely on family history), women will be encouraged to get CBEs every six to 12 months alongside annual mammography, and to consider an assortment of risk-reduction strategies as well as annual MRI. They will also be asked to begin getting these assessments at an age that is five to ten years lower than that of the youngest breast cancer case in the family.

2.4.2 Diagnosis
2.4.2.1 Morphological and Molecular Assessments

If a woman is found to possess physical symptoms of breast cancer, screening transitions into diagnosis. While breast cancer screening involves examining healthy women for signs of potential cancer, diagnostic evaluations involve further examining selected patients who are exhibiting symptoms. Post-the kind of imaging discussed above, further diagnostic imaging in the form of mammography, ultrasonography, or breast MRI may be performed to evaluate the breast. This may be coupled with chest X-rays, computerized tomography (CT) scans, bone scans, and positron emission tomography (PET) scans to assess the stage of systemic disease.
Typically, the next step following diagnostic imaging is an assessment of pathology at the molecular and morphological levels, made possible by specimen procurement and evaluation. Diseased tissue is obtained via core biopsy, fine-needle aspiration, or surgical excision for morphological evaluation\(^90\). At this level, it is important to consider that tissue preparation factors such as tissue handling, ischemic time, the use of frozen rather than fresh sections, fixation, decalcification, and processing all affect the effectiveness of evaluation\(^{91,92,93}\). Determination of gross tumor characteristics can be problematized not solely by these logistical considerations, but also by the situation of the tumor within the breast and the tumor cells’ dispersal amongst normal cells\(^94\).

Following morphological examination, immunohistochemical (IHC) and molecular tests are frequently used, particularly if morphological examination has been ambiguous\(^95\). Specific tests can be used to gain certain molecular features. For example, immunohistochemical analysis of paraffin-embedded tissue sections can determine estrogen receptor (ER), progesterone receptor (PR), and Her-2/neu (HER2) status\(^96\), and can be used to analyze RNA and DNA within these samples\(^97\). In-situ hybridization (ISH) is sometimes used to confirm initial IHC assessments of HER2 amplification\(^98\).

\subsection*{2.4.2.2 Gene Tests}

As much as the histological tumor markers isolated via these tests are of prognostic value, and are routinely used to predict patient response to therapy, they are limited in their utility by intra-tumor heterogeneity, a single tumor may possess a variety of cells, some of them positive and some of them negative for the aforementioned markers. As such, there is great interest in identifying new genes that can be included in gene tests involving DNA microarrays and high-throughput reverse transcription –
polymerase chain reaction assays. Several such genomic assays already exist, including MammaPrint [Agendia], which monitors 70 genes to predict the ten-year recurrence risk after diagnosis of stage I or stage II breast cancer regardless of hormone receptor status\textsuperscript{99,100}; Prosigna Breast Cancer Prognostic Gene Signature Assay (formerly PAM50) [NanoString], which monitors 58 genes to predict the risk of distant recurrence for post-menopausal women within ten years of diagnosis of early-stage, hormone-receptor-positive disease with up to three positive lymph nodes after five years of hormonal therapy\textsuperscript{101,102}; OncotypeDX [Genomic Health, Inc.], which monitors 21 genes to predict the risk of recurrence in early-stage, hormone-receptor-positive breast cancer as well as anticipated benefit from chemotherapy after surgery for invasive breast cancer, and the risk of recurrence, risk of ipsilateral invasive breast cancer, and anticipated benefit from radiation after surgery for ductal carcinoma \textit{in situ} (DCIS)\textsuperscript{103}; EndoPredict [Sividon Diagnostics, Myriad], which monitors 12 genes to predict the risk of distant recurrence of early-stage, hormone-receptor-positive, HER2-negative breast cancer with up to three positive lymph nodes\textsuperscript{104}; Breast Cancer Index [bioTheranostics], which monitors seven genes to predict the five- and ten-year recurrence risk of node-negative, hormone-receptor-positive breast cancer, as well as the anticipated benefit of extending hormone therapy beyond five years (for a total of ten years of hormone therapy)\textsuperscript{105}; and Mammostrat [Clarient Diagnostic Services], which monitors five genes to predict recurrence risk for early-stage, hormone-receptor-positive breast cancer\textsuperscript{106}.

2.5 Management

After a diagnosis of breast cancer is confirmed, a range of breast cancer treatments can be made available to the patient. Overall, breast cancer treatments can be
segmented into two categories: local treatments (i.e., surgery and radiation therapy), which treat the tumor without affecting the remainder of the body, and systemic treatments (i.e., chemotherapy, hormone therapy, and targeted therapy), which target cancer cells all over the body, and tend to involve the use of drugs. Many women with breast cancer will receive more than one type of treatment for their cancer.

2.6 Treatment Limitations and Future Directions
While several key treatments are available for breast cancer, breast cancer still accounts for a significant portion of female cancer deaths. Furthermore, certain key types of breast cancer, such as triple negative breast cancer (TNBC) – which accounts for 10 to 15% of all female breast cancers\(^\text{107}\) – remain associated with higher likelihood of distant recurrence and lower overall survival rates despite chemotherapy and surgery\(^\text{108}\); one study notes that no substantial progress has been made therapeutically for TNBC in over a decade\(^\text{109}\). Finally, it is important to note that even if management at first succeeds – and it often does, with systemic agents being active in 90% of primary breast cancers and 50% of metastases at the beginning of therapy – it may subsequently fail due to resistance\(^\text{110}\) (i.e., the progression of a cancer that necessitates cancer cells’ becoming immune to the therapy being administered). Resistance is not solely common – it is expected where most treatments are concerned.

As such, it is clear that there still exists a need to find more effective breast cancer treatments, particularly for certain categories of patients. For the most part, existing clinical trials involve existing drugs in novel combinations with other existing drugs, or with radiation and surgery\(^\text{111}\). Novel combinations of existing drugs aside, the eventual goal where breast cancer susceptibility gene discovery is concerned is personalized medicine\(^\text{112}\): treatment that is tailored to each individual patient, regardless of inter-
patient variability. This form of treatment is seen as having the potential to be more successful than conventional therapy because of its highly targeted approach. At present, not all patients respond well to existing cancer therapies: in fact, between 38 and 75% of all breast cancer patients fail to respond to a treatment. The principle behind personalized medicine is that if treatments were to be tailored to individual patients’ genetics, then they might better combat cancer by targeting precisely those pathways affected in a certain tumor. A key first step in this kind of personalization of treatment is the uncovering of biomarkers responsible for driving individual tumors, which include susceptibility genes, and which may be both prognostic and predictive. While next-generation technology, as has previously been mentioned, is getting increasingly competent where this is concerned, and has found significantly more biomarkers than were previously known, only 50% of hereditary breast cancers currently have known genetic bases. As such, the search for biomarkers is by no means complete.

However, even in cases where the gene or mutation responsible for a given tumor is known, personalized medicine’s success is not guaranteed. For one, treatments targeting that specific gene or mutation may not always be available. Many genes and mutations do not map directly onto a relevant existing treatment. For example, while we are aware that mutations to the PI3K-Akt pathway are causative in breast cancer, and that tumors with PIK3CA mutations may be more sensitive to PI3K inhibitors, no PI3K inhibitors are presently approved by the United States Food and Drug Administration, though several such drugs are in trials, and though inhibitors of the related mTOR pathway are available (i.e., everolimus). Similarly, while it is known that BRCA1 and BRCA2 carriers would likely benefit from poly (ADP-ribose) polymerase (PARP)
inhibitors, these drugs are not as yet definitively developed, though they are seen as likely
to be eventually successful\textsuperscript{117}.

Even when such drugs do become approved for mainstream consumption, however, personalized medicine will take some time to be adopted. Enhancing physicians’ understanding of the implications of a range of genetic tests is time-consuming, especially in light of the fact that many presently practicing physicians were trained to manage breast cancer without too many concrete genetics-based guidelines. Further, though new targeted drugs may be developed, they will be difficult to immediately disseminate, as they will need to gain not solely functional clinical trial results, but also regulatory and physician approval.
Chapter 3

Hereditary Breast Cancer

3.1 Heredity and Breast Cancer

While the majority of breast cancers are sporadic (i.e., originating from mutations acquired by the somatic cells during life), some breast cancers are hereditary (i.e., inherited by offspring from their parents via mutations in the germline DNA).

A significant proportion of cancers are known to be hereditary. A recent study showed that, of twin pairs with cancer, approximately 38% of monozygotic pairs and 26% of dizygotic pairs were diagnosed with the same type of cancer; further, monozygotic and dizygotic twins whose co-twin was diagnosed with cancer had had a 14% and 5% increased cancer risk respectively. Overall, this study found that the heritability for cancer in general was 33%, with some types of cancer – namely, prostate cancer, breast cancer, testicular cancer, kidney cancer, and melanoma – showcasing more variability due to heredity than others.

The fact that at least some proportion of breast cancers are hereditary has been understood for almost two centuries: Paul Broca identified his wife’s family as the first-ever hereditary breast cancer-prone family in 1866, and since that time, hereditary breast cancer has been well-documented. We have also known that hereditary breast and
ovarian cancer syndrome (HBOC) – i.e., BRCA-associated increased risk of female and male breast cancer, ovarian cancer, prostate cancer, pancreatic cancer, and melanoma

exists since Lynch first investigated the syndrome. Finally, the link between hereditary breast cancer and familial breast cancer has been studied for the past several decades: Ashley’s 1969 “two-hit” model of cancer susceptibility suggests that those who test positive for a certain gene or mutation – such as individuals who have relatives with hereditary cancer – are more likely to develop cancer, as they have already taken one “hit” by being related to affected individuals, and thus need only fewer excess “hits” than the average individual to develop the disease.

More recently, a study in 2000 found that breast cancer had a heritability of 27% (i.e., the proportion of variability in population disease risk due to genetic factors) among twins from Sweden, Denmark, and Finland. In 2016, following up on that 2000 study, the Nordic Twin Study of Cancer (NorTwinCan) followed up with twins from Denmark, Finland, Norway, and Sweden for an average of 32 years, and found that 31% of breast cancer risk variability may be associated with genetic factors (as well as a relatively significant estimate of shared environment influence), particularly in pre-menopausal women.

It is thought that heredity contributes in some way to 27% of breast cancers. However, only 5 to 10% of all breast cancer cases have thus far been identified with a strong inherited component, and only 4 to 5% have thus far been explained via mutations in high-penetrance, autosomal dominantly-inherited mutations. Further, despite the fact that familial breast cancer (i.e., cancer in individuals who have affected family members) may seem especially likely to have a hereditary basis, given that families possess shared genetic material, only 27% of all familial breast cancer cases – which
within themselves comprise more than 13% of all breast cancers – have been explained by way of high- or moderate-penetrance mutations; the remaining 70% of familial breast cancer cases are yet to be assigned a genetic basis, problematizing early diagnosis and treatment of high-risk individuals. It is estimated that known genetic risk loci can only explain approximately 30% of heritability\textsuperscript{129}.

This is especially troubling considering that those women who have hereditary breast cancer tend to be at a significantly higher lifetime risk of developing breast cancer (e.g., up to 80% in the case of certain mutations in \textit{BRCA1} or \textit{BRCA2}), usually develop cancer earlier in life than sporadic cases, and are likely to develop primary tumors in both breasts rather than solely one\textsuperscript{130}. Further, even women with solely a strong family history – \textit{i.e.}, women with one or more first-degree female relatives with breast cancer – and no known mutation have almost twice the risk of getting breast cancer as women without a family history\textsuperscript{131}. With more than one first-degree female relative with a history of breast cancer, the risk becomes three to four times higher; the younger the relative was upon diagnosis, too, the greater the woman’s chances of getting breast cancer. In addition to being more likely to get cancer as a whole, then, women with hereditary or familial breast cancer tend to display both earlier onset and increased severity when affected, and would thus ideally merit planning around prevention and management as early as possible.

Were all breast cancer susceptibility genes to be documented, conducting the risk estimations necessary to plan prevention and treatment – including screening, prophylactic measures, and risk management – could become considerably easier. While tracing the heredity of an individual’s cancer does not solve the problems of resistance and intra- and inter-tumor heterogeneity, or guarantee the condition’s druggability,
preventive, prognostic, and treatment-related benefits arise from understanding the genetic basis of a person’s condition. Preventively, for example, if a woman is known to have a \( BRCA1/2 \) mutation, she may have an up to 80\% life-time risk of breast cancer; annual monitoring, risk-reduction chemotherapy, and risk-reduction mastectomy may be recommended to reduce risks. Some such preventive procedures are particularly effective: namely, bilateral risk-reducing mastectomy has been shown to reduce the risk of breast cancer by up to 95\% in women with \( BRCA1/2 \) mutations, and by up to 90\% in women with a strong family history\( ^{132, 133, 134, 135} \). Prognostically, at least one-third of \( BRCA1 \) mutation carriers have triple-negative breast cancer (TNBC)\( ^{136} \), which is associated with a poorer prognosis, while \( BRCA2 \) mutations are associated with hormone receptor-positive (HR+) cancers, which tend to have better prognoses. Even in cases where no treatment exists, prognostic information may be valuable to life planning and risk management. Finally, predictively, germline mutations are not solely relevant to cancer susceptibility: they may also – like molecular subtype markers – help physicians forecast the suitability of a given treatment for a given patient based on their cancer genetics. For example, \( BRCA1 \)-mutated tumors tend to respond less well to taxanes, but may respond well to poly-ADP ribose polymerase (PARP) inhibitors and to cis-platinum\( ^{137} \).

Finally, if the genetic bases of breast cancer are not solely identified but also successfully targeted through drug development, pinpointing the genetic bases of breast cancer could lead to the advent of personalized medicine\( ^{138} \): the long-anticipated cancer care paradigm that aims to segment patients according to their individual genetics, and to utilize molecular markers to predict and address individual variability in both clinical
outcomes and drug responses. While personalized medicine has not yet been successfully implemented, the “one-dose-fits-all” approach to administering the therapies discussed in Chapter Two has been largely ineffective. Between 38 and 75 % of all patients fail to respond to a treatment. Further, of those who do respond, many experience adverse drug reactions: of all federally-approved breast cancer drugs in the United States, 16 % have been shown to cause adverse reactions, with the number of deaths due to adverse reactions exceeding 100,000 annually. With these numbers in place, as much as drug development may not always achieve as quickly or as fittingly as desired, there is a moral imperative to at least attempt to find and target the genetic bases of hereditary breast cancer.

For all these reasons, identification of relevant breast cancer susceptibility genes and especially predisposing mutations is of high significance to cancer care. Advances in genomic technologies (e.g., next-generation sequencing) are key to the complementary processes of gene discovery, cancer screening, and treatment-targeting, while functional analysis is helpful to both proposing and validating candidate susceptibility genes; these methodologies thus all also merit discussion in the context of hereditary breast cancer genetics and epidemiology, particularly in light of the fact that their use is expected to increase significantly as their costs decrease. This chapter will provide an introduction to hereditary breast cancer, known breast cancer susceptibility gene functions, and specific known genes and their clinical implications.

3.2 Understanding Hereditary Breast Cancer: Key Terminology

In order to understand hereditary breast cancer, it is important to understand certain key genetics terms that are often used to characterize hereditary cancer.
3.2.1 Mutation- and Gene-Level Terms

3.2.1.1 Mutations and Polymorphisms

As was discussed in Chapter Two, hereditary breast cancer is the result of germline mutations. By most definitions, mutations are variants that have a <1 % prevalence in the population\(^{140}\). They may be pathogenic, beneficial, or neutral. Contrastingly, polymorphisms are defined as variants that have a >1 % prevalence in the population. Polymorphisms are not typically harmful or functionally impactful in any way (although they can be), while mutations constitute changes that are often phenotypically evident. Most genetic variants are benign polymorphisms that are comparatively frequent in the population, while most pathogenic mutations are less frequent\(^{141}\).

Polymorphisms include single nucleotide polymorphisms (SNPs), and copy number variations (CNVs) (sometimes referred to as copy number alterations (CNAs) in the context of somatic tumors).

3.2.1.1.1 Single Nucleotide Polymorphisms (SNPs)

Single nucleotide polymorphisms are among the most common types of genetic variations among people, occurring once about every 300 nucleotides on average\(^{142}\), meaning that there are over 171 million SNPs\(^{143}\) in the human genome.

Typically, SNPs have no effect on health or development, though they can be utilized as markers of disease inheritance in families. However, some genes, when they possess SNPs, have been associated with breast cancer. One study has found that a polymorphism in *Sipa1* (signal-induced proliferation-associated gene 1) may have a role in germline predisposition to breast cancer metastasis, with specific SNPs within it (G313A, G2760A, and C545T) being associated with positive lymph nodes and hormone
receptor-negative disease respectively. Another member of the RAD51 family, *RAD51L1*, has also found to contain an SNP that is strongly associated with breast cancer, as have *TOX3* and *TNRC9*. Some specific SNPs have also been associated with breast cancer: rs3803662, for example, is an SNP that was found to be associated with increased breast cancer risk in Caucasian women in 2016. A total of 11 SNPs from genes like *FGFR2, MAP3K, CCND1, ZM1Z1, RAD51L11, ESR1*, and *UST* were found to be significantly associated with breast cancer in an Indian population; some of these were SNPs previously identified in a Caucasian population. Databases like the NHGRI-EBI GWAS Catalog, GWAS Central, dbSNP, and GWASdb have compiled detailed information about both very significant and less-significant genetic variants, including SNPs.

3.2.1.1.2 **Copy Number Variations (CNVs)**

Copy number variations (CNVs) are the non-mutational equivalents of large deletions and amplifications: they are changes in copy number that occur commonly enough to be thought of as polymorphic in humans$^{144}$. CNVs constitute approximately 12% of the human genome$^{145}$, and, like SNPs, can be used to trace difference between individual humans as well as disease mechanisms$^{146}$. Copy number variations are sometimes referred to as copy number aberrations when they occur in sporadic tumors$^{147}$.

Copy number variations have been known to play a role in breast cancer. In one study$^{148}$, *cis*-acting CNAs (i.e., those present on the same molecule of DNA as the gene they regulate) were associated with distinct clinical outcomes, including an estrogen receptor-positive (ER+) subgroup that nevertheless had a poor prognosis, two low-CNA subgroups that had good prognoses across luminal A, ER-positive, and ER-negative tumors, and an ERBB2-amplified subgroup with both ER-negative and ER-positive
subtypes. Another study found that germline DNA copy number variations can potentially serve as prognostic factors for breast cancer recurrence\(^{149}\): specifically, the presence of seven CNVs (two copy number gains and five copy neutral-loss of heterozygosities) was found to be associated with changes in recurrence-free survival (RFS).

### 3.2.1.2 Driver and Passenger Mutations

Most tumors contain more than one mutation: specifically, most tumors contain many passenger mutations and very few driver mutations\(^{150}\). Driver mutations are those mutations which actually cause the cancer in question, while passenger mutations are background events that have no functional significance, but nevertheless accompany drivers. Understandably, distinguishing between driver and passenger mutations within a tumor can be challenging, given that cancer drivers are uncommon and that *de novo* drivers frequently occur.

### 3.2.1.3 Genetic and Epigenetic Changes

While genetic changes are associated with breast cancer, not all inherited breast cancer is due solely to them. Epigenetic changes (*i.e.*, changes to protein expression rather than to the DNA sequence) have also been implicated in increasing breast cancer susceptibility\(^ {151}\). Two key epigenetic changes relevant to breast cancer are DNA methylation and histone modification\(^ {152}\). DNA methylation tends to silence genes that otherwise function normally. For example, hypermethylation of CpG islands (*i.e.*, CpG dinucleotide-rich regions) in the *BRCA1* promoter has been associated with breast cancer\(^ {153}\). Less commonly, breast cancer-associated promoter hypermethylation also occurs in *ESR1, HOXA5, TMS1*, and *RASSF1A*\(^ {154}\). Histone modification can involve either acetylation or methylation\(^ {155}\): increased acetylation of histones tends to result in
gene activation\textsuperscript{156}, while their methylation can sometimes result in inactivation\textsuperscript{157}.

Specific instances of histone modification have been associated with breast cancer: for example, the loss of the trimethylation of lysine 27 of H3 is associated with a negative prognosis in breast cancer\textsuperscript{158}. Aberrant hypermethylation of microRNAs is also present in primary breast tumors\textsuperscript{159}.

Specific genes are also known to be associated with epigenetic changes in breast cancer. For example, hormone receptor (ER and PR) and HER2 expression are all epigenetically regulated\textsuperscript{160}: deacetylated histones interact with the methylated ER gene promoter\textsuperscript{161}, monomethylation of K302 stabilizes ER\textsuperscript{162}, PR-negative tumors have increased PR promoter hypermethylation\textsuperscript{163}, HER2 neu-overexpressing tumors have more methylated CpG islands\textsuperscript{164}, and methylation of certain genes is correlated with HER2 amplification\textsuperscript{165}.

If epigenetic mechanisms of breast cancer susceptibility are better understood, then epigenetic biomarkers may eventually be used to develop improved breast cancer detection in blood\textsuperscript{166}, and epigenetic changes may be targeted by relevant drugs (\textit{e.g.}, methyltransferase and histone desacetylase inhibitors) in the hopes of countering epigenetic breast cancer susceptibility\textsuperscript{167}.

\textbf{3.2.1.4 Modes of Inheritance}

Breast cancer can be inherited in several different manners, depending upon which gene is causative to the inherited cancer. For the most part, however, breast cancer is inherited in an autosomal dominant manner\textsuperscript{168} (\textit{i.e.}, one mutated \textit{BRCA} allele is sufficient to make breast cancer occur): this is the case with germline pathogenic variants in \textit{BRCA1} and \textit{BRCA2}\textsuperscript{169}. Female carriers of mutations in \textit{ATM} – \textit{i.e.}, women who are
heterozygous for mutated ATM – have been shown to have a 5 to 7 times greater relative risk of breast cancer\textsuperscript{170}.

### 3.2.1.5 Penetrance

Penetrance refers to the proportion of individuals carrying a certain gene or mutation who express that gene’s or mutation’s phenotype\textsuperscript{171}. High-penetrance genes are those whose state is almost always reflected in the phenotype of carriers; intermediate-penetrance genes are those whose genotype is often but not always apparent within the phenotype; and low- or reduced-penetrance genes are those whose genotypic state is not always apparent in the phenotype\textsuperscript{172}.

Several of the most commonly alluded-to breast cancer susceptibility genes – e.g., \textit{BRCA1}, \textit{BRCA2}, \textit{TP53}, and \textit{PTEN} – are high-penetrance breast cancer genes\textsuperscript{173}. The identification of more such high-penetrance genes is critical, as the inheritance of a mutation in such a gene may increase a woman’s lifetime breast cancer risk by as much as 80\%.\textsuperscript{174} The bulk of such high-penetrance genes were discovered in the 1990s, mostly using linkage analysis and positional cloning\textsuperscript{175}. However, further attempts to find high-risk genes have largely not been fruitful: the last high-risk gene – \textit{PTEN} – was located in 1997\textsuperscript{176}, and the last moderate-penetrance gene – \textit{PALB2} – was located in 2006\textsuperscript{177}. As much as these high-risk variants are of predictive and prognostic value where they do occur, high-risk deleterious variants in genes like \textit{BRCA1}, \textit{BRCA2}, and \textit{TP53} occur in less than 0.5% of the population. As such, it has been hypothesized that familial risk is not attributable solely to rare, high-risk mutations: it can also be caused by inheritance of multiple, more common variants, each with a moderate associated increase in cancer risk\textsuperscript{178}.
This suggestion – that several variants rather than one variant can explain breast cancer risk – is known as the polygenic model of cancer susceptibility\textsuperscript{179}. The rare-variant model of cancer susceptibility suggests that overall population cancer risk is largely attributable to the joint action of several rare variants (\textit{i.e.}, SNPs with minor allele frequencies (MAFs) of $\geq 1\%$ but $\leq 5\%$, SNPs with minor allele frequencies of $\leq 1\%$, and deleterious mutations in high- and moderate-penetrance genes), with one mutation causing a disorder in a given person\textsuperscript{180}. This model argues that these variants – though they are rare – are so many in number that the disorders they cause end up being common in the population. Rare variants are typically located using candidate gene case-control studies; such rare variants have been found to lead to an up to 20-fold increased risk relative to the general population. By contrast, the common-variant model of cancer suggests that common (MAF $\geq 5\%$) SNPs jointly increase cancer risk up to three-fold compared to the general population. Common-variant searches are typically conducted using the genome-wide association study (GWAS) method\textsuperscript{181}. Specifically, regions with high levels of linkage disequilibrium (LD) (\textit{i.e.}, with several alleles that are non-randomly linked to one another at more than two loci) are located, and are then further examined for cancer susceptibility. SNPs located in breast cancer are found at a rate of 5 to 10\% in the population\textsuperscript{182}, and have been discovered to increase breast cancer risk by up to around 1.25 times relative to the general population\textsuperscript{183}.

As much as those who believe that more high-penetrance genes will be located may see finding low-penetrance variants as less useful, considering that they only cause cancer in certain cases, if we are to fully define breast cancer risk in every individual in the spirit of personalized medicine, then even low-penetrance genes will need to be
identified and targeted, as some individuals possess no mutations in high-risk genes but nevertheless present with breast cancer that appears to be hereditary. This is especially the case given that low-penetrance genes tend to be common within the population, while high-penetrance and moderate-penetrance genes tend to be rare. \textit{BRCA1} and \textit{BRCA2}, for example occur at a frequency of $< 0.5\%$ in the population\textsuperscript{184}. As such, while searching for high-penetrance genes in smaller populations remains of importance, pursuing large-scale genome-wide association studies (GWAS) in search of low-penetrance variants that cumulatively increase cancer risk is important\textsuperscript{185}.

3.2.2 Tumor-Level Terms

3.2.2.1 Intra- and Inter-Tumor Heterogeneity

Intra-tumor heterogeneity refers to the fact that cancer cells within a single tumor can differ in terms of their genetics, epigenetics, morphology, and behavior\textsuperscript{186}. According to Beca and Polyak, this is the “main obstacle to effective cancer treatment and personalized medicine”, and is prevalent across the breast cancer continuum\textsuperscript{187}. Examples of intra-tumor heterogeneity within breast cancer include differing HER2 statuses between a primary tumor and its metastasis or circulating tumor cells (CTCs), as well as invasive breast cancer tumors with some cells with HER2 amplification and some cells without it\textsuperscript{188}. Furthermore, a recent study demonstrated that driver genetic aberrations such as \textit{TP53} and \textit{PIK3CA} somatic mutations may also exhibit intra-tumor heterogeneity. Finally, tumors exhibit both spatial heterogeneity (\textit{i.e.}, differences between cells located in different parts of a given tumor) and temporal heterogeneity (\textit{i.e.}, molecular characteristics that change over time, either at the site of the original cancer or within a metastatic recurrence). All this variation even within the scope of a single tumor problematizes diagnosis and treatment: for example, a single-cell biopsy is highly
unlikely to yield genetic characteristics generalizable to the entirety of a tumor, and multiple treatments may have to be applied to the same tumor in order to successfully kill all involved cancer cells, given that different cells have different genetic characteristics.

Inter-tumor heterogeneity refers to the fact that distinct tumors within the same individual – or tumors in the same tissue across different individuals – can vary in terms of their genetics, epigenetic, morphology, and behavior. Inter-tumor heterogeneity has been somewhat well-mapped in breast cancer, and is accounted for by classifying tumors into the molecular subtypes discussed in Chapter Two.

3.3 Breast Cancer Susceptibility Genes: General Functions and Involved Pathways

Breast cancer susceptibility genes have a range of functions in their wild type state, including control of cell division, damaged or abnormal DNA recognition and response, maintenance of cell survival post-damage, DNA repair, transcription, apoptosis, senescence, autophagy, and angiogenesis. All of these functions jointly work to promote genomic stability.

3.3.1 Cell Division

In terms of controlling cell division, breast cancer susceptibility genes typically act to mediate cell division checkpoints. Prior to synthesis (S) phase, a sufficient quantity and quality must have been done during the first growth (G₁) phase to merit replication; whether or not sufficient growth has been had is monitored via the G₁/S checkpoint. Similarly, following replication and the second phase of growth (G₂), replication must have been carried out correctly and further growth must have been done in order for cell division (M phase) to occur; this is monitored by the G₂/M checkpoint. Breast cancer susceptibility genes involved in controlling these checkpoints include TP53\(^{189}\) (which
acts via p21 and cdk2 to stop cell division), CDH\textsuperscript{190}, STK11\textsuperscript{191}, ATM\textsuperscript{192}, CHEK2\textsuperscript{193}, and the MRE11-RAD50-NBN protein (MRN) complex\textsuperscript{194}.

### 3.3.2 DNA Damage Recognition and Response

Sometimes, DNA can also become damaged. This can occur for a range of reasons. Firstly, mutations in DNA polymerase\textsuperscript{195} – which typically maintains sequence fidelity by proofreading the DNA and removing mismatched bases – can inactive its proofreading function, resulting in up to a thousand-fold increase in the rate of spontaneous mutations. Secondly, both direct-acting carcinogens (\textit{i.e.}, reactive electrophiles that modify nucleotides) and indirect-acting carcinogens (\textit{i.e.}, unreactive compounds that can be converted to cancer inducers if they are not correctly metabolized and excreted by the body) can cause DNA damage\textsuperscript{196}. Many different types of DNA lesions can result\textsuperscript{197}, including a missing base, an altered base, an incorrect base, a bulge due to the deletion or insertion of a nucleotide, incorrectly linked pyrimidines, single-strand breaks (SSBs), double-strand breaks (DSBs), cross-linked strands, and 3’-deoxyribose fragments. If these lesions are left unrepaired, they can transform normal cells into rapidly proliferating cells, causing cancer. In order for these lesions to be repaired, they must first be detected; this is where some breast cancer susceptibility genes are relevant. The BRCA1-associated genome surveillance complex (BASC), for example, which includes the MRE11-RAD50-NBN protein (MRN) complex, is involved in DNA damage recognition and response\textsuperscript{198}, as are PTEN\textsuperscript{199}, ATM\textsuperscript{200}, and CHEK2\textsuperscript{201}.

### 3.3.3 DNA Repair

Repairing damaged DNA post-detection is one of the major functions of breast cancer susceptibility genes\textsuperscript{202 203 204 205 206 207}. Just as there are many forms of DNA damage, there are many mechanisms for repairing DNA damaged by chemicals,
radiation, and other carcinogens. These can be broadly categorized as mismatch repair (MMR), excision repair (spanning nucleotide excision repair and base excision repair), homologous recombination repair (HRR), and non-homologous end-joining repair (NHEJ).

### 3.3.3.1 Mismatch Repair
Mismatch repair (MMR), as its name suggests, involves identifying, removing, and replacing an incorrect nucleotide that was recently incorporated into the newly-synthesized daughter strand\(^{208}\). In other words, mismatch repair focuses on point mutations that arise from errors in replication, during genetic recombination, and via base deamination. Mismatch repair detects which strand is the parent strand via epigenetic changes: older strands of DNA are typically more methylated, and the parent strand is thus likely to contain methyl groups. Mammalian mismatch repair involves *MSH2*, *MSH6*, *MLH1*, and *PMS2*\(^{209}\). Breast cancer susceptibility genes are relevant to mismatch repair in that *BRCA1* interacts with *MSH2*.

### 3.3.3.2 Base Excision Repair (BER)
While the mismatch repair pathway only targets mismatched base pairs, base excision repair (BER) is responsible for recognizing and removing small, non-helix-distorting base lesions from the genome, and specifically on damaged bases removable via specific glycosylases\(^{210}\). These damaged bases, if not removed, could lead to mispairing or DNA breaks during replication, thus also causing mutations. The damaged bases that base excision repair targets can arise as a result of several mechanisms, including deamination, oxidation, and alkylation. Base excision repair gene defects have been shown to result in a higher mutation rate in organisms, implying that BER loss could lead to the development of cancer. In terms of breast cancer, X-ray repair cross
complementing group 1 (XRCC1), a key BER gene, has been associated with breast cancer: specifically, the XRCC1 Arg399Gln variant has been associated with increased breast cancer risk in Asians and Africans\textsuperscript{211}.

3.3.3.3 Nucleotide Excision Repair (NER)
Nucleotide excision repair (NER) specializes in recognizing and removing DNA damage induced by ultraviolet light (UV)\textsuperscript{212}. This kind of damage is characterized by bulk DNA adducts, which are mostly thymine dimers. NER involves recognizing a short, single-stranded segment of DNA that contains the lesion, removing it, and using the undamaged single-stranded DNA remains and DNA polymerase to synthesize a short complementary sequence, which is then integrated by DNA ligase. Nucleotide excision repair has been implicated in cancer, and has specifically been linked to genomic instability in breast cancer\textsuperscript{213}. Further, both BRCA1 and TP53 have been associated with nucleotide excision repair, and are prominent breast cancer susceptibility genes\textsuperscript{214}.

3.3.3.4 Homologous Recombination Repair
While MR, BER, and NER are focused mostly on repairing single-nucleotide damage, homologous recombination repair (HRR) repairs double-strand breaks (DSBs)\textsuperscript{215}. Double-strand breaks can be caused by a range of phenomena, including ionizing radiation and stalled replication forks due to lesions in the DNA or repair intermediates. If double-strand breaks are not repaired, they can, once processed, lead to mutations, loss of heterozygosity (LOH), and lethal chromosomal rearrangements, all of which can lead to cancer\textsuperscript{216}.

In general, homologous recombination repair works as follows: Initially, after a double-strand break occurs, sections of DNA around the 5’ ends of the break are removed via resection. Next, an overhanging 3’ end of the broken DNA invades a similar DNA
molecule that is not broken. Afterwards, either the double-strand break repair (DSBR) or the synthesis-dependent strand annealing (SDSA) pathway are pursued.

There are four different pathways that comprise HRR\textsuperscript{217}: these are the double-strand break repair (DSBR), synthesis-dependent strand annealing (SDSA), single-strand annealing (SSA), and break-induced replication (BIR) pathways. Different pathways are used during different periods of the cell cycle, as well as in different genomic situations; each pathway has its unique attributes. The DSBR pathway, for example, is unique in that it forms two Holliday junctions, and typically results in chromosomal crossover. The SDSA pathway results in non-crossover products, and is the major homologous recombination pathway used to repair double-strand breaks during mitosis. The SSA pathway is used to repair double-strand breaks between two repeat sequences; it is unique in that it uses the repeat sequences as its templates, and thus does not require a separate homologous molecule of DNA. Finally, the BIR pathway is used when double-strand breaks are encountered at replication forks during replication; not much is clear about its molecular mechanisms.

Homologous recombination repair genes include some of the most common breast cancer susceptibility genes (\textit{e.g.}, \textit{BRCA1}, \textit{BRCA2}, and \textit{PALB2})\textsuperscript{218, 219}.

\textbf{3.3.3.5 Non-Homologous End-Joining Repair (NHEJ)}

Finally, the non-homologous end joining (NHEJ) also targets double-strand breaks\textsuperscript{220}. However, unlike homologous recombination repair, it does not require a homologue to act as a template: instead, NHEJ just involves directly ligating the break ends\textsuperscript{221}. Typically, NHEJ uses the single-strand overhangs of the double-strand breaks to guide where it will ligate, and is usually quite precise\textsuperscript{222}. However, imprecise repair due
to incompatible overhangs can lead to loss of nucleotides, translocations, and telomere fusions, all of which can contribute to cancer\textsuperscript{223}. Further, if mutations cause the NHEJ pathway to be inactivated, then a lower-fidelity pathway known as microhomology-mediated end joining (MMEJ) takes over, and can lead to increased deletion of the DNA sequence\textsuperscript{224}. From a breast cancer perspective, \textit{BRCA1} has been implicated in non-homologous end-joining repair (NHEJ)\textsuperscript{225}.

### 3.3.4 Other Functions: Cell Survival Post-Damage, Transcription, Apoptosis, Senescence, Autophagy, and Angiogenesis

Breast cancer susceptibility genes play a range of other genomic roles. Many are involved in ensuring that the cell survives post-damage\textsuperscript{226,227,228,229}. This is primarily accomplished through ubiquitin E3 ligase activity, and is dominantly carried out by the \textit{BRCA1}- and \textit{BRCA2}-containing complex (BRCC), which includes \textit{BRCA1}, \textit{BRCA2}, \textit{BARD1}, and \textit{RAD51}\textsuperscript{230}. Some other breast cancer susceptibility genes are involved in transcription\textsuperscript{231,232}. \textit{BRCA1}, for example, interacts with RNA polymerase II\textsuperscript{233}. Others – namely, \textit{STK11} – appear to regulate \textit{TP53}-related apoptosis\textsuperscript{234}, which is critical to the adequate disposal of damaged or incorrectly proliferating cells. Some, like \textit{TP53}, regulate angiogenesis in tumors\textsuperscript{235}, which, if left unregulated, can result in tumors that grow and proliferate through increased vascularization.

### 3.3.5 Joint Functions: Gene Interactions

While certain breast cancer susceptibility genes are more directly involved in some aspects of genomic stability maintenance than others, many of the genes we will discuss in more detail later in this chapter work together. For example, \textit{BRCA1} and \textit{TP53} frequently interact, with \textit{BRCA1} regulating p53-dependent gene expression\textsuperscript{236}, and p53 modulating \textit{BRCA1} expression\textsuperscript{237}. \textit{PTEN} also regulates p53 protein levels and activity.
through several mechanisms\textsuperscript{238}, and p53, in turn, regulates PTEN\textsuperscript{239}. CHEK2 may be able to regulate p53-dependent apoptosis independently of ATM\textsuperscript{240}, but also works with ATM to regulate apoptosis; ATM, in turn, regulates TP53 through HDM2 in humans, alongside CHEK2, and directly via phosphorylation of a serine residue\textsuperscript{241}; ATM can also phosphorylate BRCA1\textsuperscript{242}, and is required for phosphorylation and activation of CHEK2 in response to infrared radiation (IR)\textsuperscript{243}. Finally, PALB2, as its name suggests, interacts directly with BRCA2 – and with BRCA1 – in homologous recombination repair\textsuperscript{244}.

\subsection{Two Cancerous Categories: Oncogenes and Tumor Suppressors}

Breast cancer susceptibility genes can be categorized into two main functional categories on the basis of their activity in their wild-type state, as well as their activity when they become mutated.

Oncogenes are genes that result in increased breast cancer susceptibility when they have unusually high function (\emph{i.e.}, if their protein expression is increased through epigenetic mechanisms, or if amplification, duplication, or some other form of gain-of-function mutation occurs at the DNA level)\textsuperscript{245}. Prior to becoming overexpressed (\emph{i.e.}, prior to becoming problematic), these genes are known as proto-oncogenes\textsuperscript{246}. Examples of oncogenes relevant to breast cancer include ERBB2\textsuperscript{247}, MYC\textsuperscript{248}, CCND1\textsuperscript{249}, and PIK3CA\textsuperscript{250}.

By contrast, tumor suppressors are genes whose normal function is to prevent tumors. They do this via a range of mechanisms, including DNA repair, cell division regulation, and programmed cell death\textsuperscript{251}. In brief, by ensuring that cells cease dividing when they need to, tumor suppressors prevent deregulation of the cell cycle (and, as a consequence, cancer). However, when tumor suppressor genes experience loss-of-function mutations, loss of heterozygosities (LOHs), or other changes resulting in
decreased function, they lose their ability to suppress tumors, and breast cancer susceptibility thus increases\textsuperscript{252}. Examples of tumor suppressor genes implicated in breast cancer include $TP53$\textsuperscript{253}, $BRCA1$\textsuperscript{254}, $BRCA2$\textsuperscript{255}, and $PTEN$\textsuperscript{256}.

3.4 Breast Cancer Susceptibility Genes: Known Mutation Types

3.4.1 Structural Types
Breast cancer is caused by a range of mutation types, which can structurally be divided into small-scale and large-scale mutations. Small-scale mutations include substitutions, insertions, and deletions.

3.4.1.1 Substitutions
Substitutions are point mutations (\textit{i.e.}, mutations that affect a single nucleotide) that involve the substitution of one base for another. Substitutions can be classified as either transitions or transversions: transitions involve a change from a purine to a purine or from a pyrimidine to a pyrimidine, and are thus typically less problematic than transversions, which involve a change from a purine to a pyrimidine or a pyrimidine to a purine, and thus alter the width of the DNA double helix\textsuperscript{257}. If a substitution results in the production of the same or a different but closely related amino acid (typically, the same amino acid), it has no significant functional consequences, and is called a silent mutation\textsuperscript{258}. Not surprisingly, silent mutations do not typically result in cancerous manifestations, as they do not alter the wild type. If a substitution results in a different amino acid than the original amino acid, it is called a missense mutation\textsuperscript{259}. Missense mutations can have significant functional consequences, given the alteration of the protein product that they imply. As such, they are quite common in breast cancer\textsuperscript{260}. Finally, if a substitution results in a stop codon and in the subsequent early truncation of a protein, then it is known as a nonsense mutation\textsuperscript{261}. This usually has significant
consequences, because the protein tends to lose function, and cellular functions are thus compromised. Substitutions differ from insertions and deletions in that they substitute one nucleotide base for another, but do not alter the overall number of bases in the DNA. If substitutions occur at two or more adjacent nucleotides, they are referred to as tandem base mutations (TBM). A large-scale survey of 560 breast cancers found twelve base substitution signatures suggesting possible driver mutations across 93 protein-coding cancer genes.

3.4.1.2 Insertions
Insertions involve the addition of one or more nucleotide base pairs into a DNA sequence. They can occur on both a small scale – e.g., with the insertion of only one incorrect base pair – and a large scale (e.g., duplication of a single exon or the entire gene). If the number of inserted nucleotides is not divisible by three, then an insertion mutation can result in a frameshift mutation (i.e., in which all the amino acids following the mutation are affected because of a changed reading frame); if the number of inserted nucleotides is divisible by three, then the mutation is said to be in-frame.

3.4.1.3 Deletions
Deletions, conversely, involve the removal of one or more nucleotide base pairs from a DNA sequence. Again, they can occur on both a small scale, involving a single base pair, and on a large scale, involving an entire piece of a chromosome. As with insertions, if deletions do not involve a multiple of three bases, then they can cause a frameshift mutation, which typically has serious consequences. Deletions, like insertions, can be either small-scale or large-scale (e.g., deletion of several exons or the entire gene).
The 560-breast-cancer survey mentioned above included six rearrangement signatures that coincided with possible driver mutations, including deletions coinciding with defective homologous-recombination-based DNA repair in BRCA1 and BRCA2.

3.4.1.4 Insertion-Deletions (Indels)
While in evolutionary biology insertions and deletions are often cumulatively called “indels”, insertion-deletions (indels) are, within the context of germline and somatic mutation biology, another class of mutations entirely. In cancer studies, an indel is a special mutation class that involves both an insertion and a deletion, typically near to one another on the chromosome, which result in a net change in the number of nucleotides. If the indel results in a net change of one to 50 nucleotides, then it is referred to as a microindel. As with insertions and deletions in general, if an indel produces a net change of a number of nucleotides that is not divisible by three, then it is known as a frameshift indel.

3.4.1.5 Large-Scale Mutations
Large-scale mutations may, for the most part, be interchangeably termed “chromosomal rearrangements”. Chromosomal rearrangements are structural changes that, unlike small-scale mutations, tend to involve the chromosome as a whole, or a large component of the chromosome. Chromosomal rearrangements include amplifications and duplications, large-scale deletions and loss of heterozygosity, inversions, and translocations.

3.4.1.5.1 Amplifications and Duplications
Amplifications are mutations that increase the number of copies of a gene in the DNA (i.e. changes in copy numbers), or that increase the quantity of RNA or DNA made from a given gene. Duplications, similarly, are mutations that increase the number of
copies of a gene in the DNA, typically doubling the prior copy number (hence “duplication”)\textsuperscript{269}. While “amplification” and “duplication” are often used interchangeably, the term “duplication” is usually used to refer to an increase in the copy number of a gene at the DNA level specifically, while the term “amplification” can be used to refer to solely protein overexpression, mediated by a change outside the DNA, and is thus not always explicitly considered a “mutation”\textsuperscript{270}. Amplifications and duplications are common in breast cancer, and are relatively well-studied, being some of the oldest mutation types discovered\textsuperscript{271}. For example, HER2 (\textit{ERBB2}) is amplified in approximately 20\% of all breast cancers\textsuperscript{272}. \textit{EGFR}, which like HER2 belongs to the human epidermal growth factor receptor family of receptor tyrosine kinase, has also been shown to be amplified and overexpressed in breast carcinoma\textsuperscript{273}.

\textbf{3.4.1.5.2 Large-Scale Deletions and Loss-of-Heterzygosities (LOHs)}

Large-scale deletions, as their name would suggest, involve deletions of massive chromosomal components, or sometimes of entire gene copies. Large-scale deletion mutations in mitochondrial DNA have been found to be associated with breast cancer\textsuperscript{274}. Similarly, loss of heterozygosity (LOH) involves the loss of one of the two original copies of a gene as well as its accompanying chromosomal region, resulting in a hemizygous – and thus inevitably “homozygous” – locus in the modified organism’s DNA. Loss of heterozygosity is often present in cancer, where it can cause an otherwise functional tumor suppressor gene to lose function; given that one entire copy of a given gene is rendered dysfunctional, LOH also often results in haploinsufficiency (\textit{i.e.}, insufficient quantities of a given protein)\textsuperscript{275}. LOH commonly occurs in \textit{BRCA1} and \textit{BRCA2} in breast cancer\textsuperscript{276}, as well as in zinc finger genes, matrix metalloproteinases, and the carcinoembryonic antigen-related cell adhesion molecule (CEACAM) family genes.
3.4.1.5.3 Inversions

Inversions are mutations in which a component of a chromosome is reversed end-to-end\(^{277}\). Inversions were found to be present in breast cancer in a study spanning 560 breast cancer whole-genome sequences\(^{278}\), in another study of \(ESR1\) mutations in hormone-resistant metastatic breast cancer\(^{279}\), and again in a third study, where inversion resulted in a frame fusion gene of \(ETV6\) and \(ITPR2\)\(^{280}\).

3.4.1.5.4 Translocations

Translocations are mutations involving chromosomal rearrangement across two non-homologous chromosomes\(^{281}\). Translocations can be reciprocal or nonreciprocal, and balanced or unbalanced\(^{282}\). If a translocation results in the joining of two otherwise-separated genes, a fusion gene (\(i.e.,\) a conjoining of the two aforementioned genes) can arise\(^{283}\). Translocations are common in breast cancer: \(NRG1\), for example, is translocated in approximately 6% of primary breast cancer cases\(^{284}\).

3.4.2 Functional Types

It is also important to be aware that mutations can be typed on the basis of their phenotypic consequences. Under this classification scheme, there are gain-of-function mutations and loss-of-function mutations.

3.4.2.1 Gain-of-Function Mutations

Gain-of-function mutations or activating mutations are mutations that introduce a new function to a given gene product, or that lead to an altered pattern of gene expression\(^{285}\). They typically act like dominant or semi-dominant alleles, and produce a novel kind of phenotype\(^{286}\). Gain-of-function mutations are also sometimes referred to as activating mutations, though the latter term can sometimes be used to refer specifically to amino acid substitutions that confer new or higher activity upon a protein. Gain-of-
function or activating mutations are prevalent in breast cancer: for example, activating mutations in HER2 have been identified even in human breast cancers that lack HER2 gene amplification\textsuperscript{287}; activating mutations have also been uncovered in \textit{ESR1}\textsuperscript{288}, \textit{PIK3CA}\textsuperscript{289}, and \textit{GATA3}\textsuperscript{290}.

\subsection*{3.4.2.2 Loss-of-Function Mutations}

By contrast, loss-of-function mutations – also called inactivating mutations – are those that result in a gene product’s having less or no function (\textit{i.e.}, being partially or wholly inactivated)\textsuperscript{291}. If a mutation results in complete loss of function (\textit{i.e.}, in a null allele), then it is called an amorphic mutation\textsuperscript{292}. Loss-of-function mutations are typically recessive\textsuperscript{293}, though some can be dominant. If loss-of-function mutations target a gene that typically acts as a tumor suppressor, then malignancy is generally promoted. Loss-of-function mutations are prevalent in many breast cancer genes, including \textit{GATA3}\textsuperscript{294}, \textit{PALB2}\textsuperscript{295}, \textit{FANCM}\textsuperscript{296}, and \textit{PTEN}\textsuperscript{297}. Importantly, \textit{PTEN} knockdown results in increased resistance to trastuzumab-based therapy\textsuperscript{298}, and can also result in hyperactivation of the PI3K pathway: specifically, activating mutations in the gene encoding the p110-alpha catalytic subunit of PI3K (\textit{PIK3CA}) have been identified in 25\% of primary breast cancers, potentially due to associated \textit{PTEN} loss\textsuperscript{299}.

\section*{3.5 Known Breast Cancer Susceptibility Genes}

\subsection*{3.3.1 Breast cancer 1 (\textit{BRCA1}) and breast cancer 2 (\textit{BRCA2})}

\subsubsection*{3.3.1.1 Discovery}
The notion that a DNA repair gene could be responsible for breast cancer susceptibility was first advanced in a paper by Hall et al. in 1990\textsuperscript{300}; this paper mapped breast cancer susceptibility to Chromosome 17q21. Four years later, Miki et al. identified breast cancer 1 (\textit{BRCA1}) at the 17q locus via positional cloning, noting its candidacy for
breast and ovarian cancer susceptibility. Breast cancer 1 thus became the first gene to be associated with hereditary breast cancer; the discovery of breast cancer 2 (BRCA2), localized to Chromosome 13q12-q13, followed within the year in a paper by Wooster et al. The discovery of these genes was anticipated to lead to increased understanding of and ability to manage breast cancer, and sparked attempts to find more genes.

3.3.1.2 Structure

BRCA1 is located on the long (q) arm of Chromosome 17, specifically at 17q21.31; it contains 81,188 base pairs, running from base pair 43,044,295 to base pair 43,125,483. It contains 24 exons, 22 of which are coding exons. In its protein form, BRCA1 is composed of 1863 amino acids. It contains four specialized domains: the Zinc finger C3HC4-type RING domain, the BRCA1 serine domain, and two BRCA1 C-terminus (BRCT) domains. The RING domain, located on the BRCA1 amino (N) terminus, interacts with BARD1 to form the BARD1 heterodimer, involved in transcriptional regulation to maintain genomic stability and DNA damage repair, including transcriptional control through modulation of chromatin structure and functioning as a corepressor. This domain is also integral to protein ubiquitination via ubiquitin E3 ligases, which is an important factor in DNA replication. The BRCA1 serine cluster domain (SCD) spans exons 11-13, and associates with BRCA2, PALB2, and ATM. Finally, the BRCT domains, located on the carboxy (C) terminus of BRCA1, associate with the histone deacetylases HDAC1 and HDAC2, and are involved in cell cycle control; they are thought to be essential to DNA repair, transcription regulation, and tumor suppression. BRCA1 also has an unusually high density of Alu repetitive DNA sequences (41.5%).
BRCA2 is located on the long (q) arm of Chromosome 13, specifically at 13q13.1.\textsuperscript{323, 324} It is sometimes also known as Fanconi anemia, complementation group D1 (\textit{FANCD1})\textsuperscript{325} because of its association with the FANC protein complex. It is 84,192 base pairs long, running from base pair 32,315,480 to base pair 32,399,672.\textsuperscript{326} It contains 27 exons.\textsuperscript{327} In its protein form, \textit{BRCA2} contains 3,418 amino acids. \textit{BRCA2} contains several special domains and motifs, including several copies of the 70-amino-acid-long BRC motif, which mediates \textit{BRCA2}-RAD51 recombinase binding and thus DNA repair.\textsuperscript{328}

\textit{BRCA1} and \textit{BRCA2} share several structural characteristics. Firstly, both have a large exon 11, and translational start sites in exon 2. Both also have AT-rich coding sequences. They also span approximately the same amount of genomic DNA (70 kb), and are both expressed at high levels in several organs (e.g., the testis). Both have also been found to contain components that biochemically and sequentially resemble members of the granin protein family,\textsuperscript{329} with \textit{BRCA1} exhibiting several characteristically granin attributes (e.g., being localized to secretory vesicles, secreted by a regulatory pathway, post-translationally glycosylated, and responsive to hormones).

3.3.1.3 Function
Both \textit{BRCA1} and \textit{BRCA2} are DNA repair genes. Specifically, both are involved in the repair of double-strand breaks (DSBs) via homologous recombination repair (HRR). Double-strand breaks are fractures of the DNA that, if inaccurately or insufficiently repaired, can lead to mutations, large-scale genomic instability, and other tumorigenic or apoptogenic changes.\textsuperscript{330} As a result, the human body has developed highly-conserved systems to detect and repair DSBs, of which \textit{BRCA1} and \textit{BRCA2} are members.\textsuperscript{331} \textit{BRCA1} is involved in both homologous recombination repair and mismatch repair (MR) through
interactions with the mismatch repair protein *MSH2*. Because they prevent tumors from developing when they are not mutated, *BRCA1* and *BRCA2* are known as tumor suppressors. They accomplish this largely by participating in an “error-free” double strand repair pathway involving *MRE11, RAD50, NBS1, ATM, CHEK2*, and *FANCJ*; when aspects of this pathway are mutated, DNA repair by lower-fidelity methods sometimes occurs, increasing the risk of cancer.

*BRCA1* and *BRCA2* accomplish their functions largely in conjunction with other genes. *BRCA1* is a member of several complexes. Firstly, it combines with several other genes (*i.e.*, *MSH2, MSH6, MLH1, ATM, BLM, PMS2*, the MRE11-RAD50-NBN protein (MRN) complex, and *RFC*) to form the *BRCA1*-associated genome surveillance complex (BASC), which may have roles in transcription-coupled repair, recognition of abnormal or damaged DNA, and DNA replication-associated repair, all of which would implicate *BRCA1* as playing a role in DNA repair and the maintenance of genomic integrity. Secondly, it acts as a component of the BRCA1-A complex, which is composed of *BRCA1, BARD1, UIMC1/RAP80, FAM175A/ABRAXAS, BRCC3/BRCC36, BRE/BRCC45*, and *BABAM1/NBA1*. Thirdly, as part of a holoenzyme complex containing *BRCA1, BRCA2, BARD1*, and *RAD51*, also known as the *BRCA1*- and *BRCA2*-containing complex (BRCC), *BRCA1* is needed for ubiquitin E3 ligase activity that enhances cellular survival following DNA damage. It also complexes with *RBBP8* to regulate *CHEK2* activation and control the G2/M checkpoint on DNA damage, and is thus also associated with regulating mitosis. It associates with a host of other genes and enzymes, including RNA polymerase II, *SMC1A*, inactive X (Xi), *COBRA1, CHEK1, CHEK2, FANCD2, FANCA, BAP1, BRCC3, AURKA, UBXN1, PCLAF,*
ACACA, and EXD2, as well as with LMO4 and CCAR2, which repress its activation. BRCA1 also induces GADD45, a DNA damage-inducible gene, and CHEK1, and is regulated by CHEK2, CTIP, and ATM, the latter two of which specifically regulate BRCA1 induction. Finally, as part of its work in homologous recombination repair (HRR), BRCA1 interacts directly with PALB2 (and, in the presence of PALB2, with BRCA2).

3.3.1.4 Breast Cancer Context: Prevalence, Mutations, Clinical Consequences, and Ethnic Distribution

When BRCA1 and BRCA2 are mutated, they tend to lose their tumor suppressor function, becoming breast cancer susceptibility genes. Jointly, they are believed to account for the largest proportion of familial breast cancer cases: up to 5% of all breast cancer cases and up to 20-25% of familial breast cancers are likely caused by mutations or rearrangements within these genes.

Breast-cancer related mutations in BRCA1 and BRCA2 vary. As early as 1994, predisposing mutations in BRCA1 including deletions, insertions, truncating mutations, and missense mutations were identified. Shortly thereafter, more than 38 common mutations were identified in a complete screen of BRCA1, with 86% of them resulting in a truncated BRCA1 protein. Today, in 2017, we are aware of more than 1600 BRCA1 mutations, most of which are frameshifts and result in either an incorrect or a dysfunctional protein. The site of a given BRCA1 mutation partially determines whether the carrier is more susceptible to ovarian cancer or breast cancer: specifically, mutant BRCA1 does not affect the growth of breast cancer cells, but mutations in the 3' portion of BRCA1 inhibited ovarian cancer cell growth. Additionally, the domain affected
matters: mutations in the 788-amino-acid RING finger domain of BRCA1 tend to be associated with familial carcinomas\textsuperscript{356}.

Like BRCA1, BRCA2 was initially linked to a relatively modest number of breast cancer-specific germline mutations (namely, six), most of which disrupted translation by leading to premature stop codons\textsuperscript{357}. At present, however, more than 1800 mutations have been found in BRCA2, spanning frameshift deletions, insertions, nonsense mutations, splicing mutations, missense mutations, and also large rearrangements of one or more exons, most of them again leading to premature termination and truncation. To this day, most BRCA2 mutations tend to be frameshift mutations\textsuperscript{358}. As with BRCA1, the region of BRCA2 affected by a given mutation matters: mutations to the “ovarian cancer cluster region” (OCCR), located in exon 11, tend to result in ovarian cancer\textsuperscript{359}, whereas certain other mutations – like the 999del5 mutation examined in an Icelandic population\textsuperscript{360} – tend to result in breast cancer risk.

Mutations in BRCA1 and BRCA2 significantly impact breast cancer risk. Specifically, women with BRCA1 or BRCA2 mutations have a lifetime risk of breast cancer of up to 80\%\textsuperscript{361}, compared to the 12\% lifetime risk that the general population faces\textsuperscript{362}. Mutations in genes within the BRCA pathway also increase the risk of certain leukemias and lymphomas up to 2000-fold\textsuperscript{363}, and are integral to BRCA1- and BRCA2-associated hereditary breast and ovarian cancer syndrome (HBOC), which is characterized by increased risk of female and male breast cancer, ovarian cancer, prostate cancer, pancreatic cancer, and, in individuals with a BRCA2 pathogenic mutation, melanoma\textsuperscript{364}. BRCA1 mutation carriers have a higher risk of associated ovarian cancer than BRCA2 mutation carriers\textsuperscript{365}. 
Each of mutated BRCA1 and BRCA2 is associated with somewhat different cancer-related presentations. For example, BRCA1 mRNA levels were associated with the transition from *in situ* to invasive breast cancer in sporadic breast cancer\(^{366}\), suggesting that BRCA1 normally negatively regulates breast epithelial cell growth. However, it isn’t solely a decrease in BRCA1 levels that results in breast cancer: aberrant location of BRCA1 rather than its absence has also been correlated with breast cancer\(^{367}\). When BRCA2 is lost, on the other hand, replicative failure, mutagen sensitivity, and genomic instability occur, creating symptoms such as those seen in Bloom syndrome and Fanconi anemia\(^{368}\).

BRCA1 and BRCA2 differ histologically in the context of breast cancers. Tumors with BRCA1 mutations tend to be triple-negative (*i.e.*, lacking the estrogen and progesterone receptors as well as HER2 overexpression), resembling basal-like and weakly luminal-like cancers, while tumors with BRCA2 tend to be hormone-receptor-positive\(^{369}\), and tend to resemble sporadic cancers\(^{370}\)\(^{371}\)\(^{372}\). A BRCA1 or BRCA2 mutation can also increase the risk of a second primary breast cancer occurrence, usually contralateral to the first: specifically, for carriers, the chance of contralateral breast cancer occurring within ten years of the first diagnosis is 10-30%, compared to the 5-10% contralateral breast cancer risk that non-carrier survivors have\(^{373}\); the lifetime risk of a second primary contralateral breast cancer for carriers is even higher, at about 40-65%\(^{374}\). Reports concerning whether the risk of ipsilateral recurrence is higher in BRCA mutation carriers are inconclusive\(^{375}\)\(^{376}\).

BRCA1- and BRCA2-mutated tumors also differ from other tumors where monitoring and treatment are concerned. For BRCA mutation carriers, annual monitoring
via mammography or MRI is recommended from age 30 onwards\textsuperscript{377}. In many cases, prophylactic surgeries, including salpingo-oophorectomy and bilateral mastectomy, are recommended to reduce mortality\textsuperscript{378} \textsuperscript{379}; specifically, oophorectomy is thought to both prevent breast and ovarian cancers in \textit{BRCA1} and \textit{BRCA2} carriers without prior breast cancer, and to reduce mortality in \textit{BRCA} mutation carriers\textsuperscript{380}. Prophylactic measures involving chemotherapy rather than surgical intervention (\textit{e.g.}, prophylactic tamoxifen) may also be administered\textsuperscript{381}. Radiation, however, may have problematic consequences: there is some evidence that radiation due to mammography before the age of 30 may increase breast cancer risk\textsuperscript{382}.

Treatment of \textit{BRCA1}-mutated tumors differs from treatment of \textit{BRCA2}-mutated tumors. Given that they are likely to be triple-negative (\textit{i.e.}, lacking the estrogen receptor, progesterone receptor, and HER2 amplification), \textit{BRCA1}-mutated tumors can be difficult to treat with conventional methods: for example, \textit{BRCA1}-mutated hormone receptor-negative tumors respond less well to taxanes (\textit{e.g.}, docetaxel, paclitaxel) than non-\textit{BRCA1}-mutated hormone receptor-negative tumors\textsuperscript{383}. However, relatively recent work has found that poly-ADP ribose polymerase (PARP) inhibitors can effectively treat \textit{BRCA1}- and \textit{BRCA2}-mutated tumors\textsuperscript{384}\textsuperscript{385}. Women with \textit{BRCA1} mutations may also benefit from olaparib or cis-platinum administration\textsuperscript{386}. By contrast, women with \textit{BRCA2} mutations primarily have ER-positive tumors\textsuperscript{387}, which can be treated with hormone therapy more readily than \textit{BRCA1}-mutated tumors.

\textit{BRCA} mutation frequencies differ from population to population. In the general population, these mutations are rare, occurring in approximately 0.1-0.5\% of the general population, translating to a 3-7\% population-attributable risk. However, certain
populations have elevated \textit{BRCA} mutation frequencies. For example, \textit{BRCA} mutations are about ten-fold more common in Ashkenazi Jews (and especially the mutations \textit{BRCA1} 185deAG, \textit{BRCA1} 5382insC, and \textit{BRCA2} 6174delT), with 2\% of this population carrying a deleterious mutation, translating to a much larger 15-30\% population-attributable risk. In the United States, \textit{BRCA1/2} mutations are especially prevalent in Ashkenazi Jews, with \textit{BRCA1} mutations presenting in 8-10\% of Ashkenazi Jewish-American breast cancer patients, and \textit{BRCA2} mutations presenting in 1\% of Ashkenazi Jewish-American breast cancer patients. Hispanic Americans, African Americans, and Caucasian Americans also have somewhat elevated mutation prevalence relative to the general population for one or both \textit{BRCA} genes: Hispanic Americans with breast cancer have a 4\% population-specific prevalence of \textit{BRCA1} mutations (data concerning \textit{BRCA2} mutations in this population is not as available), African-Americans with breast cancer have a 3\% population-specific prevalence of \textit{BRCA2} mutations, and Caucasians with breast cancer have an approximately 2-3\% population-specific prevalence of both \textit{BRCA1} and \textit{BRCA2} mutations$^{388}$. Overall, \textit{BRCA1} and \textit{BRCA2} heterozygote frequency appears to be highest in European and European American populations, followed by African-American, Latino, South Asian, East Asian, African, and European Finnish populations$^{389}$. Germline mutations of \textit{BRCA1} and \textit{BRCA2} relevant to breast cancer have also been examined in a range of smaller populations, including groups from Puerto Rico$^{390}$, the UK$^{391}$, and Finland$^{392}$.

Continued examination of individual populations in the context of germline \textit{BRCA} mutations is critical, because mutations that are recurrent in one region may not be prevalent – and thus, clinically relevant – in another. These regionally-specialized
mutations are known as founder mutations, because they tend to arise in small populations that were once “founded” primarily as a result of migration, and that are thus relatively genetically homogeneous. Examples of BRCA1 and BRCA2 founder mutations include BRCA1 185delAG and 5382insC in Hungarian and Russian Ashkenazim, BRCA2 6174delT in Hungarian and Russian Ashkenazim, BRCA1 G5193A and BRCA2 999del5 in an Icelandic population, BRCA1 IVS1-2A>G in a Finnish population, large deletions of BRCA1 in a Dutch population, BRCA1 1675delA and 1135insA in a Norwegian population, and BRCA1 2800delAA in a Scottish and Irish population.

3.3.2 Tumor protein 53 (TP53, p53)

Tumor protein 53 (TP53) is another tumor suppressor gene. Like BRCA1, it is located on chromosome 17. However, unlike BRCA1, it is located on the short (p) arm of the chromosome, specifically at 17p13.1. TP53 is 25,771 base pairs long, running from base pair 7,661,779 to base pair 7,687,550. It has 11 exons, but the first exon is non-coding. Its protein contains 393 amino acids. Like BRCA1 and BRCA2, TP53 has several highly-conserved domains that are considered essential to its function.

Functionally, TP53 is involved in the control of the G1/S and G2/M phase transitions, transcription, DNA repair, senescence, apoptosis, autophagy, mitotic catastrophe, and angiogenesis. Specifically, TP53 produces a protein that stimulates the production of p21, a protein which binds with another protein – cdk2 – to stop cell division. If TP53 is mutated, then p21 is not synthesized, and cell division cannot be stopped, which is why TP53 mutations are typically associated with cancer. TP53 also interacts with many of the same genes that BRCA1 interacts with: it induces the expression of GADD45, and is controlled by ATM/ATR and CHEK1/CHEK2.
In a somatic cancer context, p53 – the protein produced by \textit{TP53} – is among the most commonly mutated in most human cancers, with mutations occurring in about half of all cancers\textsuperscript{399}. Where breast cancer is concerned, \textit{TP53} somatic mutations are found in the majority (70-80\%) of \textit{BRCA1}-mutated cancers\textsuperscript{400}, and in approximately 20\% of breast cancer cases overall\textsuperscript{401}. Germline \textit{TP53} mutations, however, are extremely rare: while \textit{BRCA1} and \textit{BRCA2} germline mutations have an estimated population frequency of 0.5\%, \textit{TP53} germline mutations have an estimated population frequency of 0.025\%. However, when they do occur, \textit{TP53} germline mutations are associated with an 18- to 60-fold increased risk for early onset breast cancer compared to the general population\textsuperscript{402}.

As with \textit{BRCA1} and \textit{BRCA2}, the region of \textit{TP53} that is mutated matters. The majority of \textit{TP53} are missense point mutations, mostly located in the DNA binding domain; some sources have stipulated that only such mutations influence breast carcinoma outcomes at all, with mutations in other regions – like the oligomerization and transactivation domains – failing to affect disease-free and overall survival\textsuperscript{403}.

\textit{TP53} germline mutations are more common in some forms of breast cancer than in others. While the literature on \textit{TP53} germline mutation prevalence is relatively sparse, it appears that most \textit{in situ} and invasive ductal breast carcinomas with such mutations are hormone-receptor-positive (HR+) or HER2-positive (HER2+)\textsuperscript{404}. As in sporadic cancer, tumors with germline \textit{TP53} mutations also tend to cluster with \textit{BRCA1/2}-germline-mutation-associated tumors\textsuperscript{405}. Finally, sporadic \textit{TP53} mutations often occur in triple-negative breast cancers (TNBCs), which dominantly fall under the basal-like molecular class of breast cancer, and account for 15\% of all breast cancers. These cancers, as their name suggests, are negative for both hormone receptor and HER2 expression.
Characteristically, TNBCs have earlier onset and poorer outcomes, but poorer prognosis is especially like when TP53 is mutated. Overall, then, TP53 mutations tend to correlate with aggressive cancers, poor prognoses, invasion, metastasis, and expression of vascular endothelial growth factor (VEGF), which aids tumors in vascularizing – and, thus, in surviving and growing. While this does not at first seem particularly important in the context of hereditary cancer, TP53 and TNBC status become important in view of the fact that at least one third of BRCA1 mutation carriers have triple-negative breast cancer.

Certain treatments work better on TP53-mutated cancers than others, whether these cancers contain germline or somatic TP53 mutations. TP53-mutated cancers are associated with resistance to doxorubicin, anthracyclines, tamoxifen post-chemotherapy in pre-menopausal women, and radiation. Radiation should especially be avoided in patients with Li-Fraumeni syndrome, given that they are prone to have increased secondary tumor risk in response to radiation therapy. However, TP53-mutated tumors seem to respond well to paclitaxel, cyclophosphamide, and chemotherapy. TP53 status is not a predictor of the effectiveness of trastuzumab.

Relatively little has been written about TP53 germline mutations’ distribution across distinct populations. However, TP53 germline mutations are dominantly associated with one demographic: Li-Fraumeni syndrome, which in itself results in increased breast cancer risk. While germline TP53 mutations are very rare in normal breast cancer populations, existing in <1% of families, there is a very high incidence of early-onset breast cancer among Li-Fraumeni syndrome, suggesting that TP53 mutations are strong hereditary breast cancer susceptibility genes where Li-Fraumeni
syndrome patients are concerned. Germline mutations have also been isolated in certain founder populations, including in Brazil (p.R337H)\textsuperscript{417} and in Quebec (c.844C>T and c.685T>C)\textsuperscript{418}.

### 3.3.3 Phosphatase and tensin homolog (PTEN)

Phosphatase and tension homolog (\textit{PTEN}) is located on the long (q) arm of chromosome 10, specifically at 10q23.\textsuperscript{419} It spans 108,492 base pairs; it runs from base pair 87,863,438 to base pair 87,971,930. It contains nine exons\textsuperscript{420}, and 403 amino acids\textsuperscript{421}. It has several domains, including a phosphatase-coding domain in exon 5 that, when mutated, leads to a loss of tumor suppressor activity\textsuperscript{422}.

\textit{PTEN} functions to inhibit cell proliferation, survival, and growth through inactivating PI3K-dependent signaling, acting as a tumor suppressor when it is not mutated\textsuperscript{423}. \textit{PTEN} also plays a critical role in the DNA damage response, cell signaling, and genomic stability\textsuperscript{424}. \textit{PTEN} is thought to interact with \textit{BRCA1} where transcription, protein modulation, and protein stability are concerned, and to regulate \textit{TP53} levels through assorted mechanisms\textsuperscript{425}.

Certain breast cancer types are more commonly associated with \textit{PTEN} mutations than others. Specifically, the basal-like cancer subtype in hereditary \textit{BRCA1}-mutated breast cancer is more likely to see a loss of \textit{PTEN} expression\textsuperscript{426}.

\textit{PTEN} mutations are associated with Cowden syndrome, which, in female patients, is associated with a lifetime breast cancer risk of up to 50\%\textsuperscript{427}. Up to 80\% of families affected with Cowden syndrome have germline \textit{PTEN} mutations\textsuperscript{428}, making \textit{PTEN} a significant breast cancer susceptibility gene in the context of Cowden syndrome. \textit{PTEN} germline mutations have also been explored in the context of several populations. One study found that \textit{PTEN} germline mutations are unlikely to contribute to familial...
breast cancer risk in a French-Canadian population\textsuperscript{429}. Another study found that \textit{PTEN} also did not play a role in predisposing Israeli women to hereditary breast cancer\textsuperscript{430}.

\subsection*{3.3.4 Other high-penetrance genes}

While \textit{BRCA1}, \textit{BRCA2}, \textit{TP53}, and \textit{PTEN} are perhaps the most thoroughly discussed genes in the literature, there are two other high-penetrance breast cancer genes: cadherin-1 (\textit{CDH1}) and serine/threonine kinase 11 (\textit{STK11}).

Cadherin-1 (\textit{CDH1}) is located on the long (q) arm of chromosome 16, specifically at 16q22.1\textsuperscript{431}. It is 98,323 base pairs long; it runs from base pair 68,737,225 to base pair 68,835,548. It contains 16 exons\textsuperscript{432}, and 566 amino acids. \textit{CDH1} normally works to control cell division during telophase and in G1. \textit{CDH1} mutations are associated with invasive lobular breast cancer even in families without hereditary diffuse gastric cancer\textsuperscript{433}, and \textit{CDH1} mutation carriers have an up to 40-54\% lifetime risk of developing lobular breast cancer\textsuperscript{434}. As such, increased surveillance of \textit{CDH1} in at-risk populations has been recommended.

Serine/threonine kinase 11 (\textit{STK11}) is located on the short (p) arm of chromosome 19, specifically at 19p13.3. It spans 50,877 base pairs; it runs from base pair 1,177,558 to base pair 1,228,435\textsuperscript{435}. It contains nine exons encoding a 433 amino acid protein\textsuperscript{436}. \textit{STK11} is normally involved with cell cycle regulation and apoptosis control, and, like \textit{BRCA1}, \textit{BRCA2}, and \textit{TP53}, is a tumor suppressor gene. \textit{STK11} mutations are associated with Peutz-Jeghers syndrome, which is in itself associated with increased risk of cancer, including breast cancer\textsuperscript{437}. Annual breast MRI is recommended for Peutz-Jeghers patients from the age of 25-30 onwards\textsuperscript{438}. 
3.3.5 Moderate-penetrance genes

3.3.5.1 Ataxia telangiectasia mutated (ATM)

Ataxia telangiectasia mutated (ATM) is located on the long (q) arm of chromosome 11, specifically at 11q22.3439. It spans 146,648 base pairs, running from base pair 108,222,454 to base pair 108,369,102. It contains 66 exons, and 3056 amino acids440. ATM also contains a C-terminus FAT (FRAP, ATM, TRAPP) domain with a highly-conserved residue (the FATC domain), which is important for regulating ATM activity. The N-terminus domains of ATM may also be critical to interactions with substrates and other proteins.

When it functions normally, ATM is a kinase that predominantly monitors checkpoints during DNA repair, alongside BRCA1, TP53, and CHEK2. It is especially involved in the DNA damage response to DNA double-strand breaks (DSBs)441. It tends to initiate the signaling cascade containing TP53, BRCA1, and CHEK2 through phosphorylation.

Heterozygous mutations of ATM lead to a two- to five-fold increased risk of breast cancer442, depending upon the age of the carrier. Many ATM mutations are missense mutations; it has even been hypothesized that some ATM missense mutations confer higher breast cancer risk than truncating mutations. One of the most common ATM pathogenic variants is c.7271T>G 443; this mutation has been isolated in Canadian, Australian, and Northern California populations. Some studies suggest that there is an increased frequency of lobular breast cancers in ATM, while other studies dispute this. On the basis of ATM mutations’ response to PARP inhibition, it has been suggested that PARP1 inhibitors could be used to treat ATM carriers444.
3.3.5.2 Checkpoint kinase 2 (CHEK2)

Checkpoint kinase 2 (CHEK2) is located on the long (q) arm of chromosome 22, specifically at 22q12.1. It spans 94,123 base pairs, running from base pair 28,687,743 to base pair 28,741,866. It contains 14 exons, and 543 amino acids.

CHEK2 is a serine threonine kinase that primarily acts as a cell cycle checkpoint controller, and is thus part of the DNA damage response. It binds and regulates BRCA1, and, along with CHEK1, acts upstream of TP53 in the DNA damage response. It thus acts as an important conveyor of DNA damage signals to downstream repair proteins.

Two breast-cancer-linked CHEK2 mutations that are particularly prevalent in European populations are c.1100delC and I157T. The former has been shown to lead to a two-fold increased risk of breast cancer. c.1100delC is particularly common in north European populations (e.g., the Netherlands, Finland), and is associated with an increased risk of both bilateral and male breast cancer. Carrying more than one CHEK2 mutation also increases risk: in women homozygous for this mutation, breast cancer risk is six-fold compared to heterozygotes. c.1100delC is also one of three CHEK2 truncating mutations (i.e., c.1100delC, IVS2 + 1G > A, and a 5,395-base-pair founder deletion) jointly responsible for over 2% of all breast cancers in Poland. c.1100delC has important functional consequences: a truncating mutation, it induces premature termination of CHEK2’s C-terminal kinase domain, leading to decreased CHEK2 mRNA.

c.1100delC is also associated with specific forms of breast cancer: namely, its carriers tend to exhibit bilateral, estrogen-receptor-positive tumors of the ductal histological type.

I157T is a missense mutation conferring a 1.4-fold elevated risk of breast cancer, and is most common in Finland and Poland. This mutation tends to interfere
with the CHEK2 homodimerization necessary for CHEK2 activation and function. Like c.1100delC carriers, I157T carriers tend to have estrogen-receptor-positive tumors\textsuperscript{457}, though their tumors, unlike those of c.1100delC carriers, tend to be of the lobular type\textsuperscript{458}.

### 3.3.5.3 Partner and localizer of BRCA2 (PALB2)

Partner and localizer of BRCA2 (PALB2) is located on the short (p) arm of chromosome 16, specifically at 16p12.2\textsuperscript{459}. It contains 38,197 base pairs, running from base pair 23,603,160 to base pair 23,641,357. It contains 13 exons, and 1186 amino acids. It is sometimes also known as FANCN; like the BRCA genes, it is associated with the Fanconi anemia family of genes.

As its name suggests, PALB2 interacts with and manipulates BRCA2 during double-strand break (DSB) repair and homologous recombination\textsuperscript{460}. PALB2 mutations confer up to a 5.3-fold increase in breast cancer risk\textsuperscript{461}. Specifically, women with an abnormal PALB2 gene have a 35% risk of developing breast cancer by age 70\textsuperscript{462}, compared to the 80% risk associated with BRCA\textsubscript{1} and BRCA\textsubscript{2}\textsuperscript{463}. PALB2 breast-cancer-associated mutations are often loss-of-function mutations. Certain specific PALB2 founder mutations have also been located. For example, a PALB2 mutation in a Polish founder population has been estimated to potentially increase breast cancer risk\textsuperscript{464}. Polish PALB2 founder mutations are reported to be associated with lower ten-year survival compared to non-carriers (48 % versus 74.7 %)\textsuperscript{465}. Founder mutations have also been identified in Finland (c.1592delT)\textsuperscript{466} and in Canada (c.2323C>T)\textsuperscript{467}.

### 3.3.6 Plausible and former candidate genes

There are many genes that are not validated in their susceptibility status, but are being investigated. For example, RAD51C, RAD51D, XRCC2, BARD1, ABRAXAS, NBS1, RAD50, and MRE11 are all genes that are presently being investigated in terms of their
breast cancer susceptibility value. *RAD51C* and *RAD51D* seem to be mostly ovarian cancer susceptibility genes, not breast cancer susceptibility genes, but some of the other aforementioned genes merit further investigation.

One such gene is *XRCC2*, which plays a significant role in homologous recombination. Two studies identified two and ten germline *XRCC2* mutations in breast cancer families, while another study found that certain *XRCC2* SNPs – in particular, the missense mutation p.R188H – were associated with poor survival. However, *XRCC2*’s relevance to breast cancer susceptibility has been questioned, given that some studies have failed to identify mutations present in solely breast cancer patients. Another is *BRCA1*-associated RING domain (*BARD1*), which interacts with *BRCA1* during double-strand break (DSB) repair and apoptosis. *BARD1* germline mutations have been found in high-risk breast cancer families, though further, large-scale work is needed to confirm that *BARD1* may be a breast cancer susceptibility gene.

The *NBS1*, *RAD50*, and *MRE11* gene trio, responsible for the MRE11-RAD50-NBS1 (MRN) protein complex, is another group of genes involving in double-strand break (DSB) detection, DNA repair, and checkpoint control via *ATM*, *BRCA1*, and *CHEK2*. While mutations have been seen in relatively low frequencies, the *RAD50* 687delT founder deletion, the *RAD50* IVS3-1G>A splicing variant, and the *NBS1* Leu150Phe missense mutation were three mutations found in patients from Sweden, Norway, Iceland, and Finland. Finally, *ABRAXAS*, also known as *ABRA1*, *CCDC98*, and *FAM175A*, binds to *BRCA1* BRCT motifs, and acts in conjunction with *BRCA1*, *RAP80*, *BRCC36*, *BRCC45*, and *MERIT40/NBA1* to form the *BRCA1* holoenzyme complex, control DNA damage checkpoints, and respond to double-strand breaks.
(DSBs). An ABRAXAS missense mutation, p.R361Q, has been found to result in abnormal DNA responses, and has been isolated in three of 125 Finnish breast cancer families, one of 991 unselected breast cancer cases, and no healthy controls\textsuperscript{477}.

There are also genes that, once thought to increase breast cancer susceptibility, are now known not to do so. One example is BRIP1, which encodes a protein associated with BRCA1; while it was found to contribute to breast cancer susceptibility\textsuperscript{478}, more recent reports indicate that BRIP1 does not generally account for increased breast cancer risk\textsuperscript{479}, and is more an ovarian cancer susceptibility gene\textsuperscript{480, 481}.

### 3.3.7 Low-penetrant loci

While the aforementioned high- and moderate-penetrance genes account for a significant proportion of breast cancer risk, it has been estimated that a small percentage of breast cancer cases are attributable to a number of comparatively common breast cancer susceptibility loci that, operating polygenically\textsuperscript{482} or in conjunction with environmental factors, can cause elevated risk of breast cancer. These loci have been traced dominantly via genome-wide association studies (GWAS). Examples of such loci include single-nucleotide polymorphisms (SNPs) in the following genes and no-known-gene loci: MAP3K1, FGFR2, LSP1, TNRC19, TOX3, TGFB1, H19, 2q35, and 8q24\textsuperscript{483}.

**Chapter 4**
Screening *RECQL* Gene Among Breast Cancer Patients in Ontario, Canada

4.1 **RECQL: Introduction**

4.2.1 **Introduction**

RecQ-like helicase (*RECQL*) is a human gene that encodes the enzyme ATP-dependent DNA helicase Q1\(^{484}\). It is a member of the RecQ family of DNA helicases. It is also sometimes termed *RECQL1* or RecQ1. It has now been identified as a potential breast cancer susceptibility gene by both our lab\(^ {485}\) and a Chinese group\(^ {486}\).

DNA helicases, discovered in *E. coli* in 1976\(^ {487}\) and in humans in 1990\(^ {488}\), are enzymes whose main function is to separate the annealed strands of a DNA double helix by moving unidirectionally along the DNA duplex, and breaking the hydrogen bonds between annealed nucleotide bases using energy from ATP hydrolysis. Because cellular processes like replication, transcription, translation, recombination, and DNA repair all necessitate the catalysis of nucleic acid strand separation, up to one % of all eukaryotic genes are helicases\(^ {489}\).

4.2.2 **Structure**

The RecQ family of helicases is included in helicase Superfamily 2 (SF2)\(^ {490}\): the superfamily containing perhaps the largest group of helicases. Like Superfamily 1, this superfamily of helicases does not form a ring structure. It may be referred to as a beta helicase family because it works with double-stranded DNA. Even more specifically, SF2 is characterized by a specific set of motifs – Q, I, Ia, Ib, and II through VI – in its structure\(^ {491}\). Finally, RecQ helicases, as their name suggests, are structurally defined by sequences similar to those found in the catalytic domain of *E. coli* RecQ. Of the five RecQ helicases, *RECQL* is the smallest\(^ {492}\).
Mutations in certain components of RECQL can be more deleterious to gene function than mutations in other components of RECQL. Specifically, mutations found in the helicase’s highly conserved core residues as well as in the RecQ C-terminal domain of RECQL tend to affect strand separation, ATP hydrolysis, dimer formation, and protein stability more so than other mutations.

**4.2.3 Function**

RECQL belongs to a class of DExH-containing DNA helicases called the RecQ helicases. These helicases, which unwind double stranded DNA (dsDNA), are involved in important cellular functions, including DNA repair, replication, recombination, and gene transcription. Like other Superfamily 2 helicases, all five RecQ helicases help maintain genomic stability. Three of them, when mutated, are associated with genetic disorders that cause premature aging, developmental abnormalities, and cancer predisposition. For example, mutations in BLM result in Bloom syndrome (BS); mutations in RECQL4 result Rothmund-Thomson syndrome (RTS); and mutations in WRN result in Werner syndrome (WS). Incidentally, all of these syndromes are associated with a predisposition to cancer, as we suspect RECQL may also be. Furthermore, just like cells containing RECQL mutations, cells containing mutations of BLM, RECQL4, and WRN are characterized by chromosomal damage, including breaks, translocations, rearrangements, and large deletions of chromosomal components. Cells deficient in RecQ helicase products also exhibit problematic genetic recombination and replication in both human and mouse models, resulting in chromosomal instability.

While the function of RECQL has yet to be conclusively determined, several possible roles have been proposed for the gene. Perhaps most probably, RECQL may
resolve stalled replication forks, thus preventing double-strand DNA (dsDNA) breaks, maintaining genome integrity, and suppressing tumor growth.

While double-strand DNA breaks are common eukaryotic events, they must be repaired if cells are to continue functioning. Thus, two major pathways exist for repairing dsDNA breaks: homologous recombination and non-homologous DNA end joining (NHEJ). Several genes known to be breast cancer susceptibility genes are involved in repairing dsDNA breaks (e.g., BRCA1, BRCA2, PALB2)\textsuperscript{498}, suggesting that RECQL’s core proposed function (i.e., the prevention of these breaks) complements the function of known breast cancer susceptibility genes, and is thus functionally likely to also be a breast cancer susceptibility gene. RECQL may even play a role not solely in damage prevention, but also in homologous recombination repair. It is important to note that RECQL’s association with DNA repair makes it a functionally convincing candidate susceptibility gene, given that many already-known breast cancer susceptibility genes affect either the DNA damage response or some other facet of DNA repair: BRCA1 is involved in double-strand break (DSB) recognition\textsuperscript{499}; RAD50, MRE11, and NBS1 are involved in DNA end-processing\textsuperscript{500}; both BRCA genes and PALB2 are involved in strand invasion\textsuperscript{501} 502 503; RECQL appears to be involved in ATP-dependent DNA unwinding\textsuperscript{504} 505; ATM\textsuperscript{506} and CHK2\textsuperscript{507} are involved in signal transduction, and p53, the product of TP53\textsuperscript{508}, is involved in regulation of gene expression and HR branch-migration. RECQL-deficient mice, for example, are prone to aneuploidy, spontaneous chromosomal breakage, translocations, spontaneous sister chromatid exchange, increased sensitivity to ionizing radiation, a high load of dsDNA breaks, and decreased cell growth, all of which
suggest that \textit{RECQL} may normally play a role in genome integrity as a tumor suppressor, similar to the \textit{BRCA} genes\textsuperscript{509, 510}.

It is also noteworthy that \textit{RECQL} is not solely the most frequently expressed of the RecQ helicases where normal cells are concerned: it also has unusually high expression in cancer cells\textsuperscript{511}. This high RECQL expression in cancer cells seems paradoxical, given the cancer-countering functions that we attributed to RECQL above. However, as members of our lab have already suggested in a 2015 paper, it is plausible that high expression of RECQL in cancer cells only occurs after a tumor phenotype has been acquired in these cells, potentially to keep abreast of the high DNA replication rate in cancer cells, or to maintain telomere length through alternative pathways.

\textit{RECQL} interacts with several other genes, including \textit{KPNA4}\textsuperscript{512} and karyopherin alpha 2. It is also probable that \textit{RECQL} works with some of the genes listed above – \textit{i.e.}, \textit{BRCA1, BRCA2, RAD50, MRE11, NBS1, PALB2, ATM, CHK2, and TP53} – where repair is concerned. Finally, there may be a direct interaction between \textit{RECQL} and \textit{RAD51} that affects stalled replication forks.

\section*{4.3 RECQL: Breast Cancer Susceptibility Gene Candidacy in Selected Populations}

\subsection*{4.3.1 Our Lab’s Prior Paper}

Our 2015 \textit{Nature Genetics} letter identified \textit{RECQL} as a potential breast cancer susceptibility gene\textsuperscript{513}. Our lab used whole-exome sequencing to investigate a total of 195 familial breast cancer patients (both high-risk and unselected) from Polish and Quebecois populations in a discovery phase, both of which are populations with founder mutations in known susceptibility genes. We initially surveyed a range of genes in order to identify genes within which a common mutant allele was present, and then examined certain
genes in large patient sets. Genes were prioritized on the basis of possessing a truncating mutation (i.e., a nonsense mutation, essential splice-site mutation, frameshift insertion or deletion, or start codon mutation). Only variants with minor allele frequency (MAF) lesser than or equal to one % were included. While many genes harbored a mutation, only RECQL harbored multiple rare truncating mutations in both populations and was functionally similar to many existing breast cancer susceptibility genes. Specifically, two rare recurrent RECQL mutations were identified, one in each population, with each being more significantly frequent among the breast cancer patients than among the controls. Seven of 1,013 high-risk breast cancer cases compared to only one of 7,136 controls carried the c.643C>T (p.Arg215*) variant in Quebec (p < 10^{-4}), and 30 of 13,136 unselected breast cancer cases compared to only two of 4,702 controls carried the c.1667_1667+3delAGTA (p.K555delinsMYKLIHYSFR) variant in Poland (p = 0.008).

In other words, RECQL seemed to be associated with an odds ratio of 5.4 for breast cancer development among unselected Polish patients with RECQL mutations, suggesting a lifetime risk of approximately 50 %. If this risk is found to be present in further populations, then prophylactic measures like double mastectomy may be beneficial for RECQL mutation carriers. Further, one in 400 patients carried a particular RECQL truncating mutation that seemed to be particularly functionally significant. Finally, mutations increased the risk of breast cancer by sixteen-fold among higher-risk cases in Quebec. Our study also located fourteen RECQL missense mutations, all of which had a minor allele frequency lower than 0.5 %; given that rare missense mutations tend to be more likely to be pathogenic, these mutations merit further exploration where
potential pathogenicity is concerned, both via further sequencing studies and via functional assays to investigate their protein expression consequences.

However, our study involved several limitations. Firstly, it targeted two founder populations (the Quebecois population and the Ashkenazi Jewish population in Poland). Founder populations are derivative populations which arise when a group of individuals relocate to a new region and establish a community that practices some degree of endogamy, resulting in a genetically unique population that nevertheless contains a higher-than-usual frequency of certain alleles. Because such populations are prone to containing mutations distinct from those found in the general population (i.e., founder mutations), it would be wise to study a more generalized population if we wise to conclude that \textit{RECQL} mutations are relevant to other ethnicities, particularly in view of the fact that the highest incidence of breast cancer exists in countries like the United States and Canada.

Secondly, it was centered on a single recurrent truncating mutation. However, there may be other mutations in \textit{RECQL} that are pathogenic, and it is thus necessary to study the entire coding region of \textit{RECQL} if we are to truly gauge its significance as a breast cancer susceptibility gene. Finally, while our study showed that \textit{RECQL} may well be a susceptibility gene, the existence of a susceptibility gene is not in itself an actionable clinical development. In order to make \textit{RECQL} knowledge relevant to breast cancer management, it becomes important to calculate the breast cancer lifetime risk and survival associated with carrying a germline \textit{RECQL} mutation. Should the risk be found to be high in more generalized populations of breast cancer patients, recommendations
involving prophylactic procedures such as double mastectomy can subsequently be generated.

### 4.3.2 Other Papers

A 2015 paper by Sun et al. also associated mutations in *RECQL* with a predisposition to breast cancer\(^5\)\(^{14}\). In this study, 448 unrelated familial breast cancer patients were screened for *RECQL* mutations. The study found three nonsense mutations leading to a truncated protein (p.L128X, p.W172X, and p.Q266X), one mRNA splicing mutation (c.395-2A>G), and five missense mutations disrupting helicase activity (p.A195S, p.R215Q, p.R455C, p.M458K, and p.T562I). Nine of the 448 BRCA-negative families examined carried a pathogenic *RECQL* mutation compared to one of 1588 controls (\(p = 9.14 \times 10^{-6}\)), suggesting that, as per our paper, *RECQL* is a potential breast cancer susceptibility gene.

A 2016 paper by Kwong et al. again associated germline *RECQL* mutations with high-risk breast cancer in a Chinese population\(^5\)\(^{15}\). In this study, 1110 breast cancer patients negative for BRCA1, BRCA2, TP53, and PTEN gene mutations were screened for *RECQL* mutations. Four different pathogenic *RECQL* mutations were identified in six of the 1110 patients tested (0.54\%). These mutations included one frame-shift deletion (c.974_977delAAGA), two splice site mutations (c.394+1G>A, c.867+1G>T), and one nonsense mutation (c.796C>T, p.Gln266Ter). Two of these mutations were seen in more than one patient, suggesting that they may have a recurrent role in breast cancer susceptibility. Because it controls for mutations in common breast cancer susceptibility genes that we did not account for, this paper goes a step further in validating *RECQL* as a potential breast cancer susceptibility gene by eliminating the possibility that the elevated risk witnessed could be due to mutations in genes like TP53 and PTEN. The study also
underscores the significance of *RECQL* as a universal susceptibility gene by examining a population that has not as yet been considered in *RECQL* research (*i.e.*, that of Southern China).

Bogdanova et al. are another group that examined the frequency of the Polish *RECQL* mutation – c.1667_1667 + 3delAGTA – in the context of two Central European (Belarusian and German) populations. They tested 2,596 breast cancer patients and 2,132 healthy females. The c.1667_1667+3delAGTA mutation was found in 9 of 2,596 cases (0.35%) and 6 of 2,132 (0.28%) controls. Our study on the other hand, reported 30 carriers among 13,136 cases (0.23%), and 2 carriers among 4,702 controls (0.04%).

While the mutation frequency among the cases is close in both of the studies, the mutation frequency among the controls is quite high in the Bogdanova et al. study relative to what our study reported for the controls. It is also pertinent to note that the mutation frequency of this mutation in the European population as reported by the ExAC Browser after excluding the cancer genome atlas data is 0.12% (37/30,180), which is between what we reported in our study and what has been reported by Bogdanova et al. Based on the meta-analysis of the mutation frequencies among cases and controls from our study, Bogdanova et al. study and also the controls from ExAC Browser, we have 39 carriers among 15,732 cases (0.24%), and 45 carriers among 37,014 controls (0.12%) (p=1.3 x 10^{-3}). However, since ExAC Browser reports mutation frequencies from a mixture of European populations, it is still inconclusive based on the available data about the true value of the mutation frequency among the controls, and future large-scale studies from other European populations with matched cases and controls will be able to
shed more light on proving or disproving the association of RECQL
c.1667_1667+3delAGTA mutation with breast cancer risk.

In a 2017 paper, Sun et al. examined the prevalence of cancer susceptibility
gene mutations among unselected breast cancer patients in a Chinese population. 8,085
unselected Chinese breast cancer patients were enrolled, and a 62-gene panel was tested.
46 cancer susceptibility genes were found to contain germline mutations. Specifically, 30
pathogenic RECQL mutations were detected among 8,085 unselected cases.
Unfortunately, however, the Sun paper lacked controls. Looking at the 2015 Sun study’s
1,588 controls and the ExAC Browser’s 3,927 East Asian non-TCGA samples, we see
that while there were 30 mutations in 8,085 unselected cases in the 2017 Sun study, there
was only 1 carrier in 1,588 controls in the 2015 Sun study, and only 4 carriers in the
3,927 non-TCGA samples. Using these two control groups as a control for the 2017 Sun
study’s 8,085 selected cases, we can say that we would have 5 carriers in 5,515 controls
if we were to add these all to one another, and 30 carriers in 8,085 samples. Effecting a
simple case-control analysis, we see that there is a significant difference between the
number of mutations in the unselected patients and the number of mutations in the
controls (p-value: 0.0035). The associated odds ratio (OR) was 4.10.

4.4 Our Project: Screening RECQL in Ontario Population
4.4.1 Rationale
In our lab’s 2015 Nature Genetics publication, it was suggested that RECQL
mutations are associated with breast cancer susceptibility in Polish and Quebecois
founder populations. Specifically, for one recurrent truncating mutation
c.1667_1667+3delAGTA) in unselected Polish patients, an odds ratio of 5 for
developing breast cancer was observed, representing a life-time risk of about 50 %. This
suggests that \textit{RECQL} could be a high-penetrance breast cancer susceptibility gene. Like other high-penetrance breast cancer susceptibility genes, \textit{RECQL} is also rarely mutated: only about 1 in 400 cancer patients carried the aforementioned truncating mutation. But this mutation penetrance and frequency are based on a single mutation and a single population, and are thus not generalizable to the general population.

To understand whether \textit{RECQL} can be said to be an actionable susceptibility gene at the level of a heterogeneous Canadian population, we needed to arrive at a better estimate of \textit{RECQL} mutation frequency and associated breast cancer risk in the general population. We thus assessed \textit{RECQL} mutation frequency and carrier risk in a more generalized population. We also utilized a larger population, and studied mutations across the entire coding region of \textit{RECQL}: specifically, we screened the DNA of 2859 familial breast cancer patients, as well as 929 healthy controls. In order to ensure that the patients in our study were representative of hereditary breast cancer patients in Ontario, participants were recruited from Mount Sinai Hospital and Women’s College Hospital, two major genetic testing referral centers in Ontario.

Ultimately, with the data obtained from this study, we will have an improved estimate of the mutation frequency and life-time risk associated with \textit{RECQL} in a Canadian breast cancer context, which will be useful for establishing screening strategies for \textit{RECQL} in breast cancer patients, as well as prevention and treatment strategies for mutation carriers. If, for example, the 50\% life-time risk of breast cancer among \textit{RECQL} mutation carriers in our lab’s original study is consistent with the data from this research, then prophylactic measures such as double mastectomy could potentially be recommended for \textit{RECQL} mutation carriers. If, on the other hand, the life-time risk of
breast cancer associated with \textit{RECQL} mutations is not as significant in our research, then \textit{RECQL} may be found to be a breast cancer susceptibility gene solely relevant to certain founder populations rather than to a general population context: a finding that would focus future directions for \textit{RECQL} research in the contexts of the aforementioned founder populations.

4.4.2 Objective
Objective: We sought to examine the frequency of \textit{RECQL} mutations and associated life-time breast cancer risk in familial breast cancer patients and controls.

4.4.3 Methodology
Overall, we screened germline DNA for the entire coding region plus 20 base pairs of the intronic regions at each end of the exons of \textit{RECQL} in 3788 participants (2859 familial breast cancer patients and 929 controls), all of European ancestry to locate \textit{RECQL} mutation carriers. We then estimated the life-time risk of breast cancer associated with carrying a \textit{RECQL} mutation among women in Ontario.

4.4.3.1 Selection
Participants were recruited from Mount Sinai Hospital and Women’s College Hospital, two major genetic testing referral centers in Ontario. All women provided consent to participate in this study; the study protocol was approved by an ethics board. All included participants were also negative for mutations in both \textit{BRCA1} and \textit{BRCA2}. All controls were forty or older at the time of the study, and were required to have no personal history of cancer. All samples were made available in biobanks at Women’s College Hospital and Mount Sinai, where they were subsequently processed. Blood samples, medical history, and family history were collected from each participant, and several characteristics were noted, namely: current age and ethnicity. Patients recruited
from Mount Sinai will hereafter be referred to as Mount Sinai Hospital (MSH) samples, and patients recruited from Women’s College Hospital will hereafter be referred to as Women's College Hospital (WCH) samples.

4.4.3.2 Laboratory Protocol
Laboratory work was conducted at Women’s College Research Institute in Toronto, Canada.

4.4.3.2.1 Primer Design
Forty-two primer sets were designed for RECQL using the Primer3Plus program\textsuperscript{519} using the following GenBank sequences: NM_002907.3.

4.4.3.2.2 Polymerase Chain Reaction (PCR) Amplification
Testing was performed on DNA samples extracted from white blood cells. DNA was quantified using the NanoDrop ND-1000 spectrophotometer (Thermo Scientific Inc., Wilmington, DE, USA).

All 14 coding exons of RECQL (NM_002907) plus 20 base pairs from the exon boundaries were amplified in 42 amplicons between 251 and 415 in size using WaferGen SmartChip technology (WaferGen Biosystems Inc., CA, USA). First, 384-well sample plates containing the PCR reagent, genomic DNA samples, and barcoding primers were prepared. Then, assay plates containing more PCR reagent and target-specific primers were prepared. Both sets of plates were subsequently dispensed into the SmartChip using the MultiSample NanoDispenser (MSND) (WaferGen Biosystems Inc., CA, USA). Next, the SmartChip was loaded into a thermal cycler for completing the amplification and incorporating sample specific barcodes and Illumina adaptors. The amplicon library was then extracted, purified, and quantitated following the operational directions outlined in the Seq-Ready\textsuperscript{TM} TE (FLEX) manual provided by WaferGen.
4.4.3.2.3  **Sequencing**  
Sequencing was conducted at the Research Molecular Genetics Laboratory of Women’s College Research Institute (WCRI) in Toronto, Canada. Two rounds of sequencing were performed. Initially, Illumina next-generation sequencing (NGS) was performed; subsequently, Sanger sequencing was used to validate the results.

4.4.3.2.3.1  **Next-Generation Sequencing**  
Paired-end sequencing was performed using the Illumina MiSeq sequencing system (Illumina Inc., San Diego, CA, USA). Six chips were run in parallel, with each chip containing 42 samples, at 500 cycles of 250-base-pair paired-end sequencing. The mean depth of coverage was 1959x (range: 1114 to 3229). On average, 99.7% (range: 98.6% to 100%) of the coding exons of *RECQL* were covered at an 50x depth of coverage or higher (used for variant calling).

4.4.3.2.3.2  **Sanger Sequencing**  
All pathogenic variants with potential clinical significance recovered via NGS were confirmed using Sanger sequencing. Sequencing reactions were performed using a BigDye Terminator v3.1 Cycle Sequencing Kit (Fisher Scientific International Inc., Pittsburgh, PA, USA) according to the manufacturer’s protocol on the ABI prism 3500XL Genetic Analyzer (Fisher Scientific International Inc., Pittsburgh, PA, USA), and were then analyzed for variant detection using Mutation Surveyor software (SoftGenetics LLC, State College, PA, USA).

4.4.3.2.4  **Analysis**  

4.4.3.2.4.1  **General Analysis Protocol**  
Initially, the pooled sequencing reads were demultiplexed, generating two FASTQ files for each sample. Next, primer sequences were trimmed from the ends of reads, so that only the relevant DNA was preserved. Then, DNA sequences were mapped
to the hg19 – the whole human genome reference sequence – using the Burrows-Wheeler Alignment (BWA) algorithm on the Linux operating system. Following alignment, sequencing reads were stored as SAM files, which were then converted to BAM files using the Picard package. The BAM files were sorted and indexed. The files were then realigned again in order to be used in analysis of insertions/deletions (indels) and single-nucleotide polymorphisms using the Genome Analysis Toolkit (GATK). Variants were then called using UnifiedGenotyper command of the GATK. Variants were included in the analysis if they represented a nucleotide different from the RECQL reference sequence in at least 25% of reads, if they were of at least 50x depth, and if they were rare variants (minor allele frequency (MAF) ≤ 1). All truncating variants (frameshift indels, stop codon gains, essential splice site mutations) were considered pathogenic mutations, and were included. Finally, called variants were annotated using SNP & Variation Suites from Golden Helix Inc. (Bozeman, MT, USA).

4.4.3.2.4.2 **Read Mapping (Alignment) Commands**

Reads were aligned to the hg19 whole human genome reference sequence using the Burrows-Wheeler Aligner (BWA) algorithm on the Linux kernel. Our sample sequence files, which were initially in FASTQ format (i.e., the de facto standard format for storing the output of instruments like the Illumina Genome Analyzer), had to be converted to the standard Sequence Alignment/Map (SAM) format during alignment. The following command was used to convert two fastq.gz files (one for each of the two paired reads) to SAM while aligning them both to the hg19 reference sequence:

```
  bwa mem -M ucsc.hg19.fasta 1_1.fastq.gz 1_2.fastq.gz > 1.sam
```

The output of this command results in a SAM file, which can be directly manipulated using the Picard package was used to clean SAM files, and convert them to
binary versions (referred to as BAM files). BAM files were subsequently further sorted, and read groups were added. A final index of all BAM files was then generated. The following commands were used to clean, convert, and sort BAM files, and then to add read groups and build a BAM index respectively:

```
java -jar picard.jar CleanSam \
I= Sample_ID.sam \
O=Sample_ID_cleaned.sam

java -jar picard.jar SamFormatConverter \
I=Sample_ID_cleaned.sam \
O= Sample_ID_cleaned.bam

java -jar picard.jar SortSam \
I= Sample_ID_cleaned.bam \
O= Sample_ID_cleanedsorted.bam \
SORT_ORDER=coordinate

java -jar picard.jar AddOrReplaceReadGroups \
I= Sample_ID_cleanedsorted.bam \
O= Sample_ID_cleanedsortedRG.bam \
RGLB=RECQL \nRGPL=Illumina \nRGPU=hum1 \nRGSM=Sample_ID

java -jar picard.jar BuildBamIndex \
I=Sample_ID_cleanedsortedRG.bam

The indexed BAM was then subjected to further analysis via the Genome Analysis Toolkit (GATK). First, filtration was carried out to remove poor-quality reads; after filtration, reads were printed to a new BAM file. The commands used for this were as follows:

```
java -jar GenomeAnalysisTK.jar \
-T PrintReads \
-R ucsc.hg19.fasta \
-I Sample_ID_cleanedsortedRG.bam \
--read_filter MappingQualityZero \
--read_filter BadMate \
```
Once the BAM files were filtered, the RealignerTargetCreator was used to identify suspicious intervals within the sequence that were likely to be in need of realignment:

```bash
java -jar GenomeAnalysisTK.jar \
  -T RealignerTargetCreator \
  -R ucsc.hg19.fasta \
  -I Sample_ID_cleaned_sorted_Filtered.bam \
  --known 1000G_phase1.indels.hg19.vcf \
  --known Mills_and_1000G_gold_standard.indels.hg19.vcf \
  -L RECQL.intervals \
  -o 1.intervals
```

These selected intervals were then realigned using IndelRealigner, generating a BAM file of cleaned, sorted, filtered, and realigned samples.

```bash
java -jar GenomeAnalysisTK.jar \
  -T IndelRealigner \
  -R ucsc.hg19.fasta \
  -I Sample_ID_cleaned_sorted_Filtered.bam \
  -known 1000G_phase1.indels.hg19.vcf \
  -known Mills_and_1000G_gold_standard.indels.hg19.vcf \
  -targetIntervals 1.intervals \
  -o Sample_ID_cleaned_sorted_Filtered_realigned.bam \
  --consensusDeterminationModel USE_READS \
  --LOD 5.0
```

Variants were then called using UnifiedGenotyper, another Genome Analysis Toolkit feature for every 100 samples. The UnifiedGenotyper is capable of calling SNPs and insertions/deletions (indels) simultaneously, as per this command:

```bash
java -jar GenomeAnalysisTK.jar \
  -T UnifiedGenotyper \
  -R hg19.fa \
  -I Sample_ID_cleaned_sorted_Filtered_realigned.bam [-I sample2.bam ...] \
  --dbsnp_138.hg19.vcf \
  -dt NONE \
  -stand_call_conf [30.0] \
  -stand_emit_conf 10.0
```
4.4.4 Results

4.4.4.1 Overall Patient Characteristics

The mean current age of hereditary Mount Sinai Hospital breast cancer patients was 53.74 years (range: 16 – 93). The mean current age of hereditary Women’s College Hospital patients was 47.80 years (range: 23-91).

Table 1: Characteristics of Hereditary Mount Sinai Hospital Patients

<table>
<thead>
<tr>
<th>n = 2221 (Total = 2225)</th>
<th>Number</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age</td>
<td>53.74</td>
<td>N/A</td>
</tr>
<tr>
<td>Age range</td>
<td>16-93</td>
<td>N/A</td>
</tr>
<tr>
<td>Ethnicity (based on 76.13% of patients with available data)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>237</td>
<td>0.145</td>
</tr>
<tr>
<td>French</td>
<td>190</td>
<td>0.116</td>
</tr>
<tr>
<td>Italian</td>
<td>148</td>
<td>0.089</td>
</tr>
<tr>
<td>Scottish</td>
<td>119</td>
<td>0.073</td>
</tr>
<tr>
<td>Chinese</td>
<td>82</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Table 2. Characteristics of Hereditary Women’s College Hospital Patients
<table>
<thead>
<tr>
<th>n = 632 (Total = 634)</th>
<th>Number</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean age</strong></td>
<td>47.80</td>
<td></td>
</tr>
<tr>
<td><strong>Age range</strong></td>
<td>23-91</td>
<td></td>
</tr>
<tr>
<td><strong>Ethnicity</strong> (based on 98.26% of patients with available data)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>British:</td>
<td>84</td>
<td>0.133</td>
</tr>
<tr>
<td>Ashkenazi Jews</td>
<td>57</td>
<td>0.090</td>
</tr>
<tr>
<td>Chinese</td>
<td>44</td>
<td>0.070</td>
</tr>
<tr>
<td>Filipino</td>
<td>30</td>
<td>0.047</td>
</tr>
<tr>
<td>Jamaican</td>
<td>17</td>
<td>0.027</td>
</tr>
</tbody>
</table>

**4.4.4.2 Mutations and Carrier Characteristics**

Overall, 8 of 2859 hereditary breast cancer cases sampled had a mutation in *RECQL* (0.28 %), compared to only 1 of 929 controls (0.11 %) (p-value: 0.3672). One mutation (c.1476C>G, or p.Tyr492Ter) was seen twice: once in a breast cancer patient (Patient 64485) and once in a control (Patient 15385). The affected exons were 3, 6, 7, 9, 11, and 13. Six of eight breast cancer patient mutations were loss-of-function mutations; of these, one was a splice acceptor variant, two were frameshift insertion/deletions, and three were nonsense mutations (*i.e.*, gains of stop codons). The remaining two mutations were an in-frame deletion that resulted in deletion of one amino acid from the protein, and a splice region variant. The one mutation found in the control set was a loss-of-function nonsense mutation in exon 13, resulting in a gained stop codon; it involved the DNA change c.1476C>G, resulting in the protein change p.Tyr492Ter. This was the one mutation shared with one of our cases.
For the intronic mutation characterized as a splice region variant (c.867+3A>T) we ascertained its pathogenicity by utilizing standard computational tools, such as Random Forest (RF) scores, and Adaptive Boosting (AdaBoost) scores. Both scores have a range of 0 to 1, where a 0 indicates no pathogenicity, 0.6 is the threshold of possible pathogenicity, and 1 indicates a very high probability of pathogenicity\textsuperscript{520}. The intronic mutation we located had an RF value of 0.998, indicating a very high chance of pathogenicity, and an AdaBoost score of 0.748, indicating a moderately high chance of pathogenicity. Also, for the in-frame deletion (c.633_635delGAG), we found the mutation to be disruptive (affecting one nucleotide of codon #211, and two nucleotides of codon #212), causing the deletion of arginine at position 212, which is part of \textit{RECQL}’s ATP binding domain. It is also important to note that disruption of this helicase activity via a mutation within this ATP-binding domain has been documented previously. More specifically, Sun et al. found five missense mutations – one of them being p.R215Q – and confirmed the disruption of \textit{RECQL}’s helicase activity via an \textit{in vitro} helicase assay. This significantly suggests that this mutation is pathogenic.

The mean current age of breast cancer patients with 8 mutations was 46.8 years (range: 38 – 62), while the mean age of the non-carriers without a mutation was 52.45 years (range: 16-93; SD: 12.44). The difference in mean ages between carriers and non-carriers was not significant (p-value 0.1995). No ethnicity prevailed among the eight carriers with breast cancer: one patient each was Italian, Romanian, Irish, Scottish, German, South Asian and Jamaican, and one patient did not have ethnicity-related data available. Similarly, the overall cohort studied across both controls and breast cancer patients was diverse, and did not exhibit a prevailing ethnicity.
Table 3: Characteristics of *RECQL* mutations found among breast cancer patients

<table>
<thead>
<tr>
<th>Carrier ID</th>
<th>Exon</th>
<th>cDNA change</th>
<th>Protein change</th>
<th>Age</th>
<th>Gene Region (Combined)</th>
<th>Effect (Combined)</th>
<th>Sequence Ontology (Combined)</th>
</tr>
</thead>
<tbody>
<tr>
<td>321</td>
<td>6</td>
<td>NM_002907.3:c.633_635delGAG</td>
<td>NP_002898.2:p.Arg212del</td>
<td>62</td>
<td>Exon</td>
<td>Missense</td>
<td>inframe deletion</td>
</tr>
<tr>
<td>563</td>
<td>8-9</td>
<td>NM_002907.3:c.950-2A&gt;C</td>
<td>-</td>
<td>52</td>
<td>Intron</td>
<td>LoF</td>
<td>Splice acceptor variant</td>
</tr>
<tr>
<td>809</td>
<td>7</td>
<td>NM_002907.3:c.825delT</td>
<td>NP_002898.2:p.Phe275fs</td>
<td>38</td>
<td>Exon</td>
<td>LoF</td>
<td>Frameshift variant</td>
</tr>
<tr>
<td>1305</td>
<td>11</td>
<td>NM_002907.3:c.1219C&gt;T</td>
<td>NP_002898.2:p.Arg407Ter</td>
<td>-</td>
<td>Exon</td>
<td>LoF</td>
<td>Stop gained</td>
</tr>
<tr>
<td>2852</td>
<td>7-8</td>
<td>NM_002907.3:c.867+3A&gt;T</td>
<td>-</td>
<td>-</td>
<td>Intron</td>
<td>Splicing</td>
<td>Splice region variant</td>
</tr>
<tr>
<td>3022</td>
<td>9</td>
<td>NM_002907.3:c.1012_1019delGGAATTCA</td>
<td>NP_002898.2:p.Gly338fs</td>
<td>-</td>
<td>Exon</td>
<td>LoF</td>
<td>Frameshift variant</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NM_002907.3:c.1476 C&gt;G</td>
<td>NP_002898.2:p.Tyr 492Ter</td>
<td>40</td>
<td>Exon</td>
<td>LoF</td>
<td>Stop gained</td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>------------------------</td>
<td>--------------------------</td>
<td>----</td>
<td>------</td>
<td>-----</td>
<td>-------------</td>
</tr>
<tr>
<td>64485</td>
<td>13</td>
<td>NM_002907.3:c.132_1 35delGAAA</td>
<td>NP_002898.2:p.Lys 44delinsLysTer</td>
<td>42</td>
<td>Exon</td>
<td>LoF</td>
<td>Stop gained</td>
</tr>
</tbody>
</table>
4.4.4.3 Life-Time Risk Calculation

After all pathogenic truncating variants were discovered, the frequencies of pathogenic mutations in hereditary breast cancer patients and controls were calculated. Specifically, the frequency of pathogenic mutations in hereditary breast cancer patients was 0.28%; by contrast, the frequency of pathogenic mutations in controls was 0.11%.

On this basis, an odds ratio (OR) was calculated. An odds ratio is a measure of the association between an exposure (an event) and a given outcome (the presumed consequence of that event). The odds ratio represents the chances that the consequence will occur given a certain event, compared to the chances that the consequence will occur without that event, and is frequently used in case-control studies, which already inherently compare cases (i.e., outcomes that have been exposed) to controls (i.e., outcomes that have not been exposed). An odds ratio is typically calculated as such:

\[
OR = \frac{\frac{N_{\text{exposed cases}}}{N_{\text{unexposed cases}}}}{\frac{N_{\text{exposed controls}}}{N_{\text{unexposed controls}}}}
\]

where \(OR < 1\) indicates that the exposure is associated with lower odds of the outcome, \(OR = 1\) indicates that the exposure does not affect the odds of the outcome, and \(OR > 1\) indicates that the exposure is associated with higher odds of the outcome.

Specifically, in our case, the odds ratio was calculated utilizing the following equation:

\[
OR = \frac{\frac{N_{\text{RECQL mutation carriers with breast cancer}}}{N_{\text{RECQL mutation non-carriers with breast cancer}}}}{\frac{N_{\text{RECQL mutation carriers without breast cancer}}}{N_{\text{RECQL mutation non-carriers without breast cancer}}}}
\]
The OR was found to be 2.60. Based on the life-time risk associated with developing breast cancer in the general population, which is 10%, the life-time risk associated with \textit{RECQL} mutation carriers was estimated to be 26%. However, this life-time risk was insignificant, given that our p-value was 0.3672, and can thus not reliably be said to be indicative of \textit{RECQL}-associated breast cancer risk in a Canadian population.

### 4.4.5 Discussion

#### 4.4.5.1 Trends

There was no statistically significant difference between the frequency of \textit{RECQL} mutations in breast cancer patients (8/2859, 0.28%) and controls (1/929, 0.11%; p = 0.3672). Based on the odds ratio of 2.60, and 10% life-time risk associated with developing breast cancer in general population, \textit{RECQL} mutation carriers appear to have a lifetime breast cancer risk of 26%.

However, the difference between mutation frequencies in patients and controls was not as significant in this study as it was in the paper within which we initially proposed \textit{RECQL}'s breast cancer susceptibility candidacy. Furthermore, the odds ratio also decreased compared to our initial study based on Polish samples. This could due to the smaller number of cases and controls we have sequenced here. Future studies with an increased number of cases and controls will be better able to evaluate \textit{RECQL}'s role as a breast cancer susceptibility gene.

The \textit{RECQL} mutations uncovered in this study varied in terms of their characteristics. Firstly, three of the eight mutations (namely, those associated with Carrier IDs 321, 563, and 3022) do not appear to have previously been found, while the other five had previously been described within either our 2015 study or online resources (\textit{i.e.},
the ExAC Browser, Reference SNP, and Nature’s full mutation list). Mutations which have not previously been charted do not often have well-documented functional consequences; as such, depending upon their eventual functional consequences, they could be either significant or insignificant to RECQL’s breast cancer susceptibility gene potential. Secondly, the eight mutations had distinct effects: six were loss-of-function (LoF) mutations, one was a in-frame deletion, and one was a splicing mutation; specifically, in terms of sequence ontology, three involved the gaining of a stop codon, two were splice region or acceptor variants, two were frameshift variants, and one was an in-frame deletion. Finally, while six of the eight mutations were exonic, two were intronic.

In summary, our results suggest a statistically non-significant association of RECQL with breast cancer. Several large studies are under way by our team and others; their results could help clarify the association of RECQL with breast cancer. Until more information is gathered, nothing definite can be said in the way of lifetime risk; for that reason, no prophylactic or national screening measures can be recommended.

4.4.5.2 Limitations

There are two major limitations in this study. Firstly, due to lack of large number of cases and controls, we were unable to validate the role of RECQL as a potential breast cancer susceptibility gene. Instead, we were only able to screen the entire coding regions of the gene, and report the relevant mutation frequency. In order to make a definitive statement on the role of RECQL in breast cancer, we need to conduct large-scale case-control studies.
Secondly, and more importantly, in our current study, we only looked at the role of truncating variants in RECQL. We found 13 missense mutations, but we did not investigate their roles further. However, RECQL can only thoroughly be investigated if the missense variants from this study and the original study by Cybulski et al. are analyzed. There are several indications that missense mutations may be critical to future studies involving RECQL. First, the study conducted by Sun et al. in 2015, which independently found RECQL to be a breast cancer susceptibility gene, found nine mutations. The majority of them – five – were missense, and these missense mutations were confirmed to be pathogenic via an in-vitro functional assay. Second, Sun et al. conducted a follow-up study in 2017 examining 8085 unselected breast cancer patients which retrieved 30 mutations in RECQL. Once again, 20 out of the 30 mutations were missense mutations that had previously been confirmed to be pathogenic. Third and most recently, Tervasmäki et al. found a rare founder missense variant in a Finnish population, and no mutations in their controls. The mutation was found in the helicase domain of RECQL, which was also found to be disrupted by an in-frame deletion mutation in our study, and a missense mutation in the 2015 Sun et al. study. Fourth, in the discovery paper of RECQL by Cybulski et al., there were 14 missense variants found in 1,145 patients, and all of them had a minor allele frequency (MAF) of less than 1%. Though no functional analysis has thus far been done on these missense variants, the fact that these variants have an MAF of less than 1% indicates a strong likelihood of pathogenicity. Fifth, there are several other genes in which the role of missense mutations has been studied extensively. These include TP53, ATM, and CHEK2, all of which are breast cancer susceptibility genes, and are involved in some facet of DNA repair or DNA
damage signaling response. Finally, given that RECQL is a DNA helicase, it should not be surprising to see that missense mutations might play a larger role than truncating variants, as has often been the case with other breast cancer susceptibility genes. The beta hairpin secondary structure of RECQL, which is used for unwinding double-stranded DNA, has in fact been shown to be easily disrupted by both a missense variant (found in Sun et al. study), and an in-frame indel (found in our lab’s original RECQL discovery study, Cybulski et al.). Due to the six reasons mentioned above, it seems quite plausible that by conducting in-vitro functional assays for the 13 missense variants from our current study, and 14 missense variants from the Cybulski et al. study, we will find a clearer picture of the role of RECQL in breast cancer susceptibility.

Given that missense mutations may play a much stronger role than truncating mutations, it is premature to make a definitive statement about the role of RECQL in breast cancer on the basis of the results and odds ratio found in our current study. Not only is the number of cases and controls lower, but in failing to consider the role of missense mutations, we may thus far be unaware of the mutation frequency and general hereditary mutation spectrum of RECQL. Further, the findings of in-vitro functional assays in follow-up studies will shed more light on RECQL mutations in a heterogeneous population like that of Ontario.
Chapter 5

Old Genes, New Context: BRCA1, BRCA2, and PALB2 in the Caribbean

5.1 Cancer in Selected Populations

In the RECQL-related component of our laboratory work, we focused on screening a breast cancer susceptibility gene initially discovered in two founder populations in a generalized, heterogeneous population context.

The same is true of populations in developing countries. The problem of cancer in developing countries was recently designated a high-priority issue for the International Atomic Energy Agency (IAEA), which has launched a Joint Program on Cancer Control in collaboration with the World Health Organization (WHO). According to 2010 WHO statistics, approximately 70% of the world’s cancer deaths are currently occurring in the developing world, in large part because 70% of cancer patients in developing countries are diagnosed at a very late stage of illness, when treatment is no longer effective. As of 2010, approximately 30 developing countries did not even possess a single radiation therapy machine. Thus, cancers that are highly curable in some parts of the world (e.g., cervical cancer) have a survival rate of approximately 20% in certain developing regions. In view of the fact that one in six cancers worldwide are caused by infections, and most infection-attributable cancers occur in less developing countries, the fact that the burden of cancer-related health problems has shifted to less developed countries while the gap in cancer-related health outcomes between developing and developed countries is widening is a critical problem.
5.2 Breast Cancer in the Caribbean

The Caribbean is classified as a developing region by the United Nations\textsuperscript{521}, and is heavily burdened by both cancer in general, with cancer being the second leading cause of death in the Caribbean\textsuperscript{522}, and by breast cancer specifically, with breast cancer being the leading cause of cancer-related death among Caribbean women\textsuperscript{523}. In a study conducted by the Caribbean Public Health Agency (CARPHA) and the United States Center for Disease Control and Prevention (CDC), it was discovered that the rates of death from cervical, breast, prostate, and colon cancers are two to nine times higher in the Caribbean than in the United States\textsuperscript{524}. Sixty-three % of Caribbean women with breast cancer are also pre-menopausal\textsuperscript{525}, indicating increased prevalence of early-onset cancers in the region. Further, an alarming percentage of women present with late-stage cancers, indicating that early diagnosis, prevention, and treatment are likely not being practiced.

The prevalence of breast cancer death in the Caribbean is largely due to situational factors. Epidemiologically, the Caribbean is presently transitioning from a developing- to a more developed-country in terms of fertility pattern\textsuperscript{526}. Because of this, the age at menarche is decreasing, age at menopause is increasing, the number of children women have is decreasing, and the age at first pregnancy is increasing; all of these factors modify breast cancer risk, as is further discussed in Chapter Two. Furthermore, breast cancer prevention and management is not well-orchestrated in the Caribbean, and awareness is not normalized: tests for ER and HER2 are not carried out in many countries, screening mammography is often not feasible, and breast self-examination is culturally stigmatized. This means that late-stage breast cancer predominates, as self-examination – the diagnostic method responsible for detecting 86 % of breast cancers in Caribbean women – is only carried out when tumors are noticeable\textsuperscript{527}. 
One of the important components of ameliorating screening, prevention, and treatment in the Caribbean is understanding the breast cancer genetic landscape to see if existing genetic testing can theoretically be implemented, and if screening recommendations can be drafted. To this end, we undertook two projects focused on Caribbean populations.

5.3 Project One: A survey of BRCA1, BRCA2, and PALB2 mutations in women with breast cancer in Trinidad and Tobago

5.3.1 Rationale
Trinidad and Tobago has one of the highest breast cancer incidence rates in the Caribbean \(^{528}\), but the country’s hereditary breast cancer burden has not been studied \(^{529}\). While access to screening and treatment is free, breast cancer is often not diagnosed at advanced stages \(^{530}531\ 532\ 533\) , indicating the possibility that hereditary breast cancer susceptibility is not adequately pre-empted. Island nations also often have unique genetic characteristics due to their immigration patterns and reproductive isolation, and our laboratory has already verified the contribution of mutations in \(BRCA1/2\) to breast cancer in the Bahamas, another Caribbean nation \(^{534}\). \(BRCA1/2\) also tend to have a higher mutation rate in non-Caucasian populations, and Trinidad and Tobago’s population is largely (approximately 80 %) African and Indian. \(BRCA1/2\) are also highly penetrant genes, with mutated genes conferring up to an 80 % life-time risk of breast cancer \(^{535}\). Mutations in moderate penetrance genes, such as \(PALB2\), also result in a moderate-to-high (33 to 58 %) life-time risk of breast cancer \(^{536}\). If the \(BRCA\) genes and \(PALB2\) are found to be significantly prevalent in an unselected population, then offering genetic testing for these three genes to all breast cancer patients in Trinidad and Tobago could be merited.
5.3.2 Objectives
We sought to examine the frequency of \textit{BRCA1/2} and \textit{PALB2} mutations in unselected Trinidad and Tobago breast cancer cases to determine whether genetic screening for all three genes for all breast cancer patients should be implemented in Trinidad and Tobago.

5.3.3 Methods
Participants were recruited from the largest breast cancer referral center in Port of Spain, Trinidad and Tobago, and in the breast clinic of the St. James Medical Complex, which is a part of the North West Regional Health Authority in Trinidad and Tobago. All women consented to participate in this study in writing, and the University of Miami Institutional Review Board (IRB) and the ethics committee of the North West Regional Health Authority of Trinidad and Tobago approved our protocol. Women were eligible to participate if they had been diagnosed with invasive primary breast cancer or ductal carcinoma \textit{in situ} (DCIS) at any age, in any year, and if at least one of their grandparents was born in Trinidad and Tobago. A total of 268 women participated. Saliva samples, medical history, and family history were collected, reviewed, and subsequently verified via medical record reviews or surgical pathology reports if available. Several breast cancer-related characteristics – namely, age at diagnosis, laterality, stage, grade, histology, estrogen receptor (ER) status, HER2/neu (HER2) status, age at menarche, age at first pregnancy, parity, birth order, race, ethnic group, and age at menopause – were also collected.

Saliva was collected using the Oragene \textregistered DNA sample collection kit (OG-250 format, DNA Genotek, Kanata, ON, Canada), and extracted following the manufacturer’s
instructions. DNA was then quantified using the NanoDrop ND-1000 Spectrophotometer (Thermo Scientific Inc., Wilmington, DE, USA).

All 22 coding exons of \textit{BRCA1} (NM_007294.3), 26 coding exons of \textit{BRCA2} (NM_000059.3), and the first 13 coding exons of \textit{PALB2} (NM_024675.3) plus 20 base pairs from the exon boundaries were amplified in 81, 134, and 54 amplicons respectively using Wafergen SmartChip technology (Wafergene Inc., CA, USA). Following incorporation of adaptors and sample-specific DNA barcodes via PCR, the prepared DNA library was pooled and paired-end-sequenced at 2 x 250 cycles using an Illumina MiSeq sequencer.

Reads were aligned to the reference sequences of the three genes using the Burrows-Wheeler Aligner (BWA) algorithm, and subsequently converted from SAM file format to BAM files, which could then be sorted and indexed. Reads that remained unmapped, exhibited low mapping quality, or mapped to more than one region were filtered using the Genome Analysis Tool Kit (GATK). SNPs, insertions, and deletions were subsequently called using the GATK’s UnifiedGenotyper module. In order to be utilized in calling variants, regions had to have at least 20-fold depth of coverage; in order to be considered a variant, a nucleotide had to differ from the reference sequence in at least 25 % of reads aligned to a given position. After variants were called, they were annotated using the SNP and Variation Suite (GoldenHelix Inc., Bozeman, MT, USA).

Sanger sequencing was then used to validate all truncating and missense variants of clinical significance; specifically, sequencing reactions were performed using the BigDye Terminator v3.1 Cycle Sequencing Kit (Fisher Scientific International Inc., Pittsburgh, PA, USA) according to the manufacturer’s protocol on the ABI prism
3500XL Genetic Analyzer (Fisher Scientific International Inc., Pittsburgh, PA, USA), and were analyzed for variant detection using Mutation Surveyor (SoftGenetics LLC, State College, PA, USA).

The mean depth of coverage of all samples was 1927x (range: 907-4967x). On average, 95.2 % (range: 87.4-100 %) of the coding exons were covered at a 20x depth of coverage or higher.

5.3.4 Results
The mean age of diagnosis was 42.9 years (range: 19-80). The majority of patients (84.7 %) were pre-menopausal at diagnosis. Breast cancer was also self-detected in the majority of patients (91.6 %). Eleven patients had bilateral breast cancer (4.1 %). 115 of 231 patients (49.8 %) had ER-positive breast cancers, 42 of 223 patients (18.8 %) had HER2-positive cancers, and 91 of 235 patients (38.7 %) had ER- or HER-negative cancers.

28 of 268 patients (10.4 %) had a mutation in BRCA1, BRCA2, or PALB2. Of these individuals, 15 (53.6 %) had BRCA1 mutations, 10 (35.7 %) had BRCA2 mutations, 2 (7.1 %) had PALB2 mutations, and 1 (3.6 %) had both a BRCA2 and a PALB2 mutation. Of the 29 mutations seen in total, four mutations were seen twice in unrelated patients.

The mean age at diagnosis of the 28 mutation carriers was 39.8 years (range: 20-48), and the mean age of diagnosis of the 239 non-carriers was 43.3 years (range: 19-80; p-value: 0.052). The mean age at diagnosis was 37.3 years for BRCA1 mutation carriers, 42.2 years for BRCA2 mutation carriers, 43.3 years for non-carriers, and 46 years for PALB2 mutation carriers. The prevalence of mutations was 33.3 % (4/12) in women diagnosed between the ages of 20 and 29, 8.4 % (7/83) in women diagnosed between the ages of 30 and 39, 13.2 % (18/136) in women diagnosed between the ages of 40 and 49,
and 0 % (0/35) in women diagnosed at age 50 or older. For women aged 49 or younger, the rate of mutation was 12.5 % (29/231).

None of the eight participants with DCIS carried a mutation. Mutations were found in 17.6 % (16/91) of ER-negative, HER2-negative breast cancers, and in 6.1 % (7/115) ER-positive breast cancers. None of the BRCA1/2 carriers were HER2-positive, but one PALB2 carrier was HER2-positive.

The most commonly-reported ethnicities were African, Indian, and Caucasian. African heritage was higher in carriers (85.7 %, 24/28) than in non-carriers (57.1 %, 137/240; p-value: 0.003).

17.2 % (46/268) of patients had a first-degree breast cancer relative, and 34 % (91/268) of patients had a first- or second-degree breast cancer relative. 12.3 % (33/268) patients had a first- or second-degree ovarian cancer relative. The prevalence of mutations was 31.8 % (7/22) among women with two or more first- or second-degree breast cancer relatives, 14.9 % (10/67) among women with one first- or second-degree breast cancer relatives, and 6.4 % (11/172) among women with no first- or second-degree breast cancer relatives. 73.3 % of BRCA1 mutation carriers and 50 % of BRCA2 mutation carriers had a first- or second-degree breast cancer relative compared to 30.8 % of non-carriers (p-value: 0.001). 64.3 % (18/28) carriers had a first- or second-degree history of breast or ovarian cancer compared to 38.3 % of non-carriers (p-value: 0.008).
Table 4. Characteristics of the 28 Trinidadian breast cancer patients who carry a pathogenic mutation in BRCA1, BRCA2 and PALB2 genes

<table>
<thead>
<tr>
<th>Carrier ID</th>
<th>Gene</th>
<th>Exon</th>
<th>cDNA change</th>
<th>Protein change</th>
<th>Age at Dx</th>
<th>Stage</th>
<th>Grade</th>
<th>Hormone status</th>
<th>Ethnicity</th>
<th>Family history - 1st or 2nd degree relatives with breast or ovarian cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNT007</td>
<td>BRCA2</td>
<td>11</td>
<td>c.6373_6375delA CCinsG</td>
<td>44</td>
<td>I</td>
<td>II</td>
<td></td>
<td>ER+, HER-</td>
<td>African, Indian, Caucasian.</td>
<td>No</td>
</tr>
<tr>
<td>TNT012</td>
<td>BRCA1</td>
<td>11</td>
<td>c.3108delT</td>
<td>42</td>
<td>IIa</td>
<td>II</td>
<td></td>
<td>ER-, HER-</td>
<td>Indian</td>
<td>No</td>
</tr>
<tr>
<td>TNT013</td>
<td>BRCA1</td>
<td>11</td>
<td>c.1636_1654delA TGAATATTACT AATAGTG</td>
<td>28</td>
<td>III</td>
<td>III</td>
<td></td>
<td>ER-, HER-</td>
<td>African</td>
<td>Yes</td>
</tr>
<tr>
<td>TNT021</td>
<td>BRCA1</td>
<td>21</td>
<td>c.5324T&gt;G</td>
<td>20</td>
<td>T3N0M0</td>
<td>III</td>
<td></td>
<td>ER-, HER-</td>
<td>African</td>
<td>No</td>
</tr>
<tr>
<td>ID</td>
<td>Gene</td>
<td>Age</td>
<td>Mutation</td>
<td>Protein</td>
<td>Age</td>
<td>Tumor Stage</td>
<td>Tumor Node</td>
<td>Tumor Metastasis</td>
<td>Race</td>
<td>Ethnicity</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>-----</td>
<td>------------------</td>
<td>---------</td>
<td>-----</td>
<td>-------------</td>
<td>------------</td>
<td>-----------------</td>
<td>--------</td>
<td>-----------------</td>
</tr>
<tr>
<td>TNT023</td>
<td>BRCA1</td>
<td>24</td>
<td>c.5510G&gt;A</td>
<td>p.Trp1837Ter</td>
<td>37</td>
<td>T2N0M0</td>
<td>N/A</td>
<td>Unknown Unknown</td>
<td>Chinese, Caucasian, African.</td>
<td>Yes</td>
</tr>
<tr>
<td>TNT026</td>
<td>PALB2</td>
<td>4</td>
<td>c.1571C&gt;G</td>
<td>p.Ser524Ter</td>
<td>43</td>
<td>T2N0M0</td>
<td>N/A</td>
<td>ER-HER+</td>
<td>African Syrian</td>
<td>Yes</td>
</tr>
<tr>
<td>TNT027</td>
<td>BRCA1</td>
<td>11</td>
<td>c.3331_3334delC</td>
<td>AAG</td>
<td>42</td>
<td>IVII</td>
<td>II</td>
<td>ER-HER-</td>
<td>Mixed</td>
<td>Yes</td>
</tr>
<tr>
<td>TNT032</td>
<td>BRCA1</td>
<td>13</td>
<td>c.4327C&gt;T</td>
<td>p.Arg1443Ter</td>
<td>48</td>
<td>?? N/A</td>
<td>ER-HER-</td>
<td>African, Caucasian.</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>TNT033</td>
<td>BRCA1</td>
<td>11</td>
<td>c.2389_2390delG</td>
<td>A</td>
<td>44</td>
<td>T2N0M0</td>
<td>N/A</td>
<td>ER-HER-</td>
<td>African, Indian</td>
<td>No</td>
</tr>
<tr>
<td>TNT034</td>
<td>BRCA2</td>
<td>10</td>
<td>c.1763_1766delA</td>
<td>TAA</td>
<td>48</td>
<td>T3N1M0</td>
<td>N/A</td>
<td>ER-HER-</td>
<td>African</td>
<td>Yes</td>
</tr>
<tr>
<td>Patient ID</td>
<td>Gene</td>
<td>Chromosome</td>
<td>Mutation Description</td>
<td>Age at Diagnosis</td>
<td>T Stage</td>
<td>N Stage</td>
<td>M Stage</td>
<td>ER Status</td>
<td>Her2 Status</td>
<td>Ethnicity</td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td>------------</td>
<td>----------------------</td>
<td>-----------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>-----------</td>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td>TNT036</td>
<td>BRCA1</td>
<td>19</td>
<td>c.5177_5180delGAAA</td>
<td>44</td>
<td>T2N0M0</td>
<td>N/A</td>
<td>ER-HER-</td>
<td>African</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>TNT038</td>
<td>BRCA2</td>
<td>10</td>
<td>c.1308_1309delGA</td>
<td>41</td>
<td>T3N1M0</td>
<td>III</td>
<td>ER-HER-</td>
<td>African, Indian</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>TNT042</td>
<td>BRCA2</td>
<td>25</td>
<td>c.9382C&gt;T p.Arg3128Ter</td>
<td>37</td>
<td>T2N0M0</td>
<td>N/A</td>
<td>Unknown Unknown</td>
<td>African</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>TNT046</td>
<td>BRCA1</td>
<td>11</td>
<td>c.2766delA</td>
<td>44</td>
<td>T3N0M0</td>
<td>III</td>
<td>ER-HER-</td>
<td>Indian</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>TNT067</td>
<td>BRCA2</td>
<td>11</td>
<td>c.5909C&gt;A p.Ser1970Ter</td>
<td>39</td>
<td>T1N2M0</td>
<td>III</td>
<td>ER+HER-</td>
<td>Chinese, African, Caucasian, Caribbean</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>TNT076</td>
<td>BRCA2</td>
<td>15</td>
<td>c.7558C&gt;T p.Arg2520Ter</td>
<td>38</td>
<td>T4dN2M0</td>
<td>III</td>
<td>ER+HER-</td>
<td>African, Caucasian, Caribbean</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ID</td>
<td>Gene</td>
<td>ID</td>
<td>Variant</td>
<td>Description</td>
<td>Tumor Size</td>
<td>Stage</td>
<td>Node</td>
<td>Metastasis</td>
<td>Hormone Status</td>
<td>Tumor Description</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>------------</td>
<td>-------</td>
<td>------</td>
<td>------------</td>
<td>-----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>TNT084</td>
<td>BRCA2</td>
<td>18</td>
<td>c.7977-1G&gt;A</td>
<td></td>
<td>47</td>
<td>II</td>
<td>T1N0 M0</td>
<td></td>
<td>ER-Her-</td>
<td>African, Caucasian</td>
</tr>
<tr>
<td>TNT088</td>
<td>BRCA1</td>
<td>11</td>
<td>c.2138C&gt;A</td>
<td>p.Ser713Ter</td>
<td>40</td>
<td>III</td>
<td>T2N1 M0</td>
<td></td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>TNT130</td>
<td>BRCA1</td>
<td>11</td>
<td>c.2686_2687insA</td>
<td></td>
<td>26</td>
<td>III</td>
<td>T2N0 M0</td>
<td></td>
<td>ER-Her-</td>
<td>Indian, African</td>
</tr>
<tr>
<td>TNT132</td>
<td>PALB2</td>
<td>4</td>
<td>c.1571C&gt;G</td>
<td>p.Ser524Ter</td>
<td>47</td>
<td>N/A</td>
<td>T2N0 M0</td>
<td></td>
<td>ER+Her-</td>
<td>African, Indian, Caucasian</td>
</tr>
<tr>
<td>TNT141</td>
<td>BRCA2</td>
<td>11</td>
<td>c.6437_6440delA</td>
<td>TCA</td>
<td>48</td>
<td>N/A</td>
<td>T2N3 M0</td>
<td></td>
<td>ER-Her-</td>
<td>Indian, African, Caucasian</td>
</tr>
<tr>
<td>TNT141</td>
<td>PALB2</td>
<td>1</td>
<td>c.12_13insT</td>
<td></td>
<td>48</td>
<td>N/A</td>
<td>T2N0 M0</td>
<td></td>
<td>ER-Her-</td>
<td>Indian, African, Caucasian</td>
</tr>
<tr>
<td>TNT147</td>
<td>BRCA2</td>
<td>25</td>
<td>c.9382C&gt;T</td>
<td>p.Arg3128Ter</td>
<td>35</td>
<td>III</td>
<td>T3N1 M0</td>
<td></td>
<td>ER-Her-</td>
<td>African</td>
</tr>
<tr>
<td>TNT151</td>
<td>BRCA1</td>
<td>11</td>
<td>c.3365_3366delCA</td>
<td></td>
<td>28</td>
<td>III</td>
<td>T2N1 M0</td>
<td></td>
<td>ER+Unknown</td>
<td>African</td>
</tr>
<tr>
<td>Case</td>
<td>Gene</td>
<td>Exon</td>
<td>Mutation</td>
<td>Age</td>
<td>Tumor Stage</td>
<td>ER Status</td>
<td>HER Status</td>
<td>Ethnicity</td>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>------</td>
<td>----------</td>
<td>-----</td>
<td>-------------</td>
<td>-----------</td>
<td>------------</td>
<td>-----------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>TNT190</td>
<td>BRCA2</td>
<td>10</td>
<td>c.1310_1313delAAGA</td>
<td>40</td>
<td>T2N2M0</td>
<td>II</td>
<td>ER+HER-</td>
<td>African</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>TNT241</td>
<td>BRCA1</td>
<td>16</td>
<td>c.4945_4947delAGAinsTTTTTT</td>
<td>33</td>
<td>T3N1M0</td>
<td>III</td>
<td>ER-HER-</td>
<td>African</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>TNT249</td>
<td>BRCA2</td>
<td>15</td>
<td>c.7558C&gt;Tp.Arg2520Ter</td>
<td>47</td>
<td>III</td>
<td>N/A</td>
<td>ER+HER-</td>
<td>IndianSpanish</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>TNT253</td>
<td>BRCA1</td>
<td>11</td>
<td>c.3764_3765insA</td>
<td>48</td>
<td>T2NxM0</td>
<td>III</td>
<td>ER-HER-</td>
<td>IndianAfrican</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>TNT267</td>
<td>BRCA1</td>
<td>16</td>
<td>c.4945_4947delAGAinsTTTTTT</td>
<td>36</td>
<td>T2N0M0</td>
<td>N/A</td>
<td>ER-HER-</td>
<td>CaribbeanSpanishIndian</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>
5.3.5 Discussion
The demographic insights we derived from this study are of relevance. Trinidad and Tobago has an ethnically diverse population due to its colonial history: its published ethnic composition is 39.5% African, 40.3% East Indian, 18.4% mixed, 1.2% Caucasian, and 0.6% Chinese, along with a small percentage of other groups. The distribution of ethnic groups self-reported in our study was similar to this composition, suggesting that our study results could be generalizable. Notably, as in past studies, a significant number of mutation carriers reported either solely African ancestry (10/28, 35.7%) or African and mixed African ancestry (24/28, 85.7%). Several of our mutations had previously been found in our other Caribbean cohorts in the Bahamas and Barbados, and reported in Afro-Caribbean, African-American, and Nigerian breast cancer patients previously. This suggests that African ancestry is inherently connected to certain BRCA1/2 mutations, knowledge of which could ameliorate targeted screening programs. Twelve mutation carriers also reported either Indian-only or mixed Indian ancestry, and two of the mutations found in our Indian patients had previously been reported in an East-Indian patient and two Muslim families in Israel, suggesting that these two mutations may be founder mutations associated with Indian ancestry.

Where carrier status is concerned, 52.7% of patients diagnosed at age 45 or under tested negative for mutations, but 36% of mutation-negative patients reported a family history of first- or second-degree breast or ovarian cancer. This suggests that there may be other genes which account for the susceptibility of these additional cases; Churpek et al. found that 20% of mutations in a high-risk African-American breast cancer population were attributable to PALB2, CHEK2, BARD1, ATM, PTEN, and TP53, and Desmond later found that an mutation detection increased by 3.8% when these additional
genes were added\textsuperscript{543}. However, the interpretability of low- and moderate-penetrance gene results in multigene panels is uncertain.

Two other factors may have influenced the generalizability of our study. Firstly, the average age of our participants was 43, while the average age of women diagnosed with breast cancer in Trinidad and Tobago is considerably higher, at 56.5\textsuperscript{544}. Self-selection may have resulted in a younger population’s being overrepresented. Secondly, we enrolled women in the North West Regional Health District, which has the highest incidence of breast cancer and the highest mortality rate in Trinidad and Tobago, which could have resulted in somewhat poorly generalizable results.

Finally, in the way of overarching insights, this is the first survey of hereditary breast cancer in Trinidad and Tobago, indicating a need for further research into breast cancer susceptibility and genetic testing. Overall, we detected \textit{BRCA1}/\textit{2} mutations in 9.5\% of cases, and \textit{PALB2} mutations in 1.1\% of cases, and showed that mutations in \textit{BRCA1}, \textit{BRCA2}, and \textit{PALB2} account for more than 10\% of breast cancer in Trinidad and Tobago. This suggests that a national screening program would potentially aid in guiding prevention and therapy.

\textbf{5.4 Project Two: A high frequency of PALB2 mutations in Jamaican patients with breast cancer}

\textbf{5.4.1 Rationale}

Jamaica has one of the highest breast cancer incidence rates in the Caribbean\textsuperscript{545}, but the country’s breast cancer genetic landscape has not been explored, and while genetic testing and counselling is offered to women with a family history of breast and ovarian cancer, unselected women are not yet encouraged to undergo testing. Further, while \textit{BRCA1} and \textit{BRCA2} are known contributors to hereditary breast cancer in Jamaica,
PALB2 has not yet been examined as a potential addition to genetic testing panels. If the BRCA genes and PALB2 are found to be significantly prevalent in an unselected population, then instituting a formal national screening policy, as has been advocated for by several prior studies, could be merited, and could foster cancer risk reduction and survival improvements.

5.4.2 Objectives
We sought to examine the frequency of BRCA1/2 mutations in unselected Jamaican breast cancer cases to determine whether genetic screening for all breast cancer patients should be implemented in Jamaica. We also examined the frequency of PALB2 mutations in unselected patients in order to determine whether PALB2 should potentially be added to genetic testing panels.

5.4.3 Methods
Participants were recruited from the University of West Indies, MONA, Jamaica. All women consented to participate in this study in writing, and the UHWI/UWI/FMS Ethics Committee of the University of the West Indies approved our protocol. Women were eligible to participate if they had at least one Jamaican-born grandparent. A total of 179 women participated. Saliva samples, medical history, and family history were collected and reviewed, and several breast cancer-related characteristics – namely, age at diagnosis, presence of second primary cancers, tumor size, nodal status, histology, grade, estrogen receptor status, and HER2 receptor status – were noted.

The remainder of our methodology for this study is identical to the methodology employed in the previous Trinidad and Tobago study – for a more detailed overview of our methods, please refer to Section 5.3.3.
The mean depth of coverage of all samples was 2948x (range: 1055-6550x). On average, 98.2 % (range: 92.4-100 %) of the coding exons were covered at a 20x depth of coverage or higher.

5.4.4 Results

The mean age of diagnosis was 49 years (range: 26-76). 94 of 177 patients (53.1 %) were pre-menopausal at diagnosis. 91 of 141 patients (64.5 %) had ER-positive breast cancers, 33 of 136 patients (24.3 %) had HER2-positive cancers, and 38 of 146 (26 %) had ER- or HER-negative cancers.

82 of 175 patients (46.9 %) reported a female relative with breast cancer, and 12 of 177 (6.8 %) reported a female relative with ovarian cancer.

8 of 179 patients (4.5 %) had a mutation in BRCA1, BRCA2, or PALB2. Of these individuals, 1 (12.5 %) had a BRCA1 mutation, 2 (25 %) had BRCA2 mutations, and 5 (62.5 %) had PALB2 mutations. No mutation was seen in more than one patient.

The mean age at diagnosis of the eight mutation carriers was 50 (range: 41-60), which was similar to the mean age of diagnosis of the 171 non-carriers (49 years; range: 26-76; p-value: 0.790). The prevalence of mutations was 0 % (0/4) in women diagnosed at age 39 or younger, 6 % (3/50) in women diagnosed between the ages of 40 and 49, and 5.7 % (5/88) in women diagnosed at age 50 or older.

A mutation was found in 0 % (0/38) ER-negative and HER2-negative breast cancers, and in 7.7 % (7/91) of ER-positive breast cancers. There were no HER2-positive carriers.

The prevalence of mutations was 8.5 % (7/82) in women with one or more breast cancer relatives, and 1.1 % (1/93) in women with no breast cancer relatives (p-value: 0.03). 87.5 % (7/8) carriers had a family history of breast or ovarian cancer compared to
46.8 % (80/171) non-carriers (p-value: 0.02). All three \textit{BRCA1} and \textit{BRCA2} carriers had a breast cancer relative, while only three of five (60 \%) of \textit{PALB2} carriers had a breast cancer relative. No \textit{BRCA1} or \textit{BRCA2} carriers had an ovarian cancer relative, but one of five (20 \%) of \textit{PALB2} carriers had an ovarian cancer relative.
Table 5. Characteristics of the 8 Jamaican breast cancer patients who carry a pathogenic mutation in BRCA1, BRCA2 and PALB2 genes.

<table>
<thead>
<tr>
<th>Carrier ID</th>
<th>Gene</th>
<th>Exon</th>
<th>cDNA change</th>
<th>Protein change</th>
<th>Age at Dx</th>
<th>Stage</th>
<th>Grade</th>
<th>Hormone status</th>
<th>Ethnicity</th>
<th>Family history (1+ relatives with breast or ovarian cancer)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JAM 039</td>
<td>BRCA1</td>
<td>10</td>
<td>c.1953_1956 delGAAA</td>
<td>p.Lys653fs</td>
<td>47</td>
<td>-</td>
<td>-</td>
<td>ER + HER2 -</td>
<td>African</td>
<td>Yes</td>
</tr>
<tr>
<td>JAM 051</td>
<td>BRCA2</td>
<td>14</td>
<td>c.7386_7389 delCAAT</td>
<td>p.Asn2463fs</td>
<td>46</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>African</td>
<td>Yes</td>
</tr>
<tr>
<td>JAM 074</td>
<td>PALB2</td>
<td>5</td>
<td>c.2052delC</td>
<td>p.Arg686fs</td>
<td>52</td>
<td>-</td>
<td>-</td>
<td>ER + HER2 -</td>
<td>African, Chinese, Indian</td>
<td>Yes</td>
</tr>
<tr>
<td>JAM 089</td>
<td>PALB2</td>
<td>3</td>
<td>c.109C&gt;A</td>
<td>Probably affecting splicing*</td>
<td>60</td>
<td>-</td>
<td>-</td>
<td>ER +</td>
<td>HER2 -</td>
<td>African</td>
</tr>
<tr>
<td>JAM 139</td>
<td>PALB2</td>
<td>1</td>
<td>c.43G&gt;T</td>
<td>p.Glu15Ter</td>
<td>41</td>
<td>II</td>
<td>-</td>
<td>ER +</td>
<td>HER2 -</td>
<td>Native American, African</td>
</tr>
<tr>
<td>JAM 150</td>
<td>BRCA2</td>
<td>15</td>
<td>c.7558C&gt;T</td>
<td>p.Arg2520Ter</td>
<td>54</td>
<td>IV</td>
<td>-</td>
<td>ER +</td>
<td>HER2 -</td>
<td>African</td>
</tr>
<tr>
<td>JAM 156</td>
<td>PALB2</td>
<td>11</td>
<td>c.3166C&gt;T</td>
<td>p.Gln1056Ter</td>
<td>52</td>
<td>IIIC</td>
<td>-</td>
<td>ER +</td>
<td>HER2 -</td>
<td>-</td>
</tr>
<tr>
<td>JAM 177</td>
<td>PALB2</td>
<td>4</td>
<td>c.758_759insT</td>
<td>p.Ser254fs</td>
<td>50</td>
<td>-</td>
<td>-</td>
<td>ER +</td>
<td>-</td>
<td>African, Indian</td>
</tr>
</tbody>
</table>

* This is a nucleotide change located at the 5’ end of exon 3 of PALB2 gene. This mutation is predicted to affect the splicing of the PALB2 mRNA according to dbSNV database with an Ada and RF scores of 0.93 and 0.70 respectively. 

108
5.4.5 Discussion
Where BRCA1/2 mutations are concerned, our study’s mutation rate (1.7 %) was lower than expected. In most Caucasian non-founder populations, the prevalence of BRCA mutations in unselected patients is higher (3-4 %) than the frequency we found. In Trinidad and Tobago and the Bahamas, two countries with African ancestry comparable to Jamaica’s, we identified considerably higher BRCA1/2 mutation rates at 9.5 % and 24 % respectively; a previous study of unselected patients also found an 11 % BRCA1/2 mutation rate in Nigeria, another country with predominantly African ancestry. In general, this suggests that the BRCA1/2 mutation rate of Jamaica may not have been thoroughly detected because of the size of our sample, and that further studies examining BRCA1/2 mutations in the context of the Jamaican population may be merited.

In terms of PALB2, we found a relatively high frequency of mutation carriers among Jamaican breast cancer patients (5/179, 2.8 % compared to 3/179, 1.7 % for BRCA1/2). This frequency of PALB2 mutations among unselected breast cancer patients is the highest to be reported to date: in Finland, a single relevant PALB2 mutation was identified in 1 % of unselected breast cancer patients, and in Trinidad and Tobago, 1.1 % (3/268) of unselected breast cancer patients had a PALB2 mutation. Further, PALB2 mutation carriers have a significant life-time risk of breast cancer (between 33 and 58 %). Our study results also appeared to be generalizable to the broader Jamaican population, which is approximately 90.9 % African, as 99 % of our patients reported at least some African ancestry. In light of all this, PALB2 may be a gene that is particularly relevant to Jamaican breast cancer susceptibility. Incorporation of PALB2 into Jamaican gene panels could potentially improve mutation detection rates in Jamaica.
Finally, as has already been noted, Jamaica has a high prevalence of breast cancer, but relatively few studies examining breast cancer in Jamaica specifically have been conducted, and though several studies have called for urgent institution of a formal national screening policy since 2008, no such policy exists. As such, we would encourage the development of governmental programs geared at promoting genetic counseling, testing, and the investigation of further risk management and testing for at-risk family members. Such programs, if implemented, could potentially ameliorate risk-reducing technique implementation, leading to cancer risk reduction and survival improvement.
Chapter 6

Future Directions

6.1 RECQL

*RECQL* was first proposed as a potential breast cancer susceptibility gene by our team, and independently by a Chinese group, in 2015. Since the work previously done was localized to founder populations, our team initially decided to validate these findings in a more heterogeneous population. However, since the number of cases in this project was not sufficient to perform validation, we resorted to screening *RECQL* in a more heterogeneous population.

Our lab’s next step would be to further expand our understanding of *RECQL* using genetic epidemiology and molecular genetics approaches. More specifically, our lab intends to first determine the frequency of pathogenic *RECQL* mutations, including truncating and missense mutations, among 6000 hereditary breast cancer patients, unselected breast cancer patients, and healthy women in non-founder populations. Our lab’s goal from this work would be to build on the work of this current project, and to validate *RECQL*’s role as a potential breast cancer susceptibility gene.

Next, based on the missense mutations obtained (both from our original study, and our follow-up studies), our lab intends to perform functional assays to ascertain the pathogenicity of *RECQL* missense mutations. Our lab will specifically look for DNA damage foci formation, and DNA repair pathways *in vivo*. As part of this project, our lab and our collaborators will also knock-down expression of *RECQL* using RNA interference techniques such as small interfering RNA (siRNA) and short hairpin RNA (shRNA). Our lab will conduct assays to monitor genomic instability and to see if missense mutations affect H2AX and 53BP1 levels. The two major DNA repair pathways
– homologous recombination repair and non-homologous end-joining repair – will be assayed in vivo to study the role of RECQL in DNA repair. Finally, our lab will also test how RECQL missense mutations affect cell survival post-damage in order to understand whether RECQL mutations are resistant in the presence of genotoxic environments.

Additionally, in order to establish a causal link between RECQL mutations and breast cancer, our lab will conduct knock-out (KO) mice studies looking at wild-type RECQL heterozygous (+/-), and RECQL homozygous (-/-) mice in the presence of mammary carcinogen DMBA (7, 12-dimethylbenz(alpha)anthracene) to observe mammary tumor incidence. These studies will help our lab answer the question of whether RECQL knock-out (KO) mice have a higher risk of developing mammary tumors compared to wild-type mice, demonstrating RECQL’s pathogenicity in vivo.

Finally, our lab will perform gene expression profiling to identify genes that are specifically upregulated or downregulated in RECQL-deficient tumour cells, and to measure expression levels of RECQL in RECQL mutation carriers and non-carriers among tumour cells in order to better understand the genetic signatures associated with RECQL mutation carriers for screening purposes. To do this, our lab will conduct whole-mRNA sequencing of breast tumour cells in 40 patients with germline RECQL mutations and 40 controls. These controls will be matched for histology, grade, and hormone status [ER, PR and HER2] of tumours, and for menopause status of the patients. Patients who have received neoadjuvant therapy will be excluded. Knowing the particular gene expression patterns will ameliorate the development of targeted treatments for breast cancer patients with germline RECQL mutations.
6.2 BRCA1, BRCA2, & PALB2 in the Caribbean

Overall, it is clear that hereditary breast cancer is a significant concern across the Caribbean, and merits implementation of national screening policies. However, there are also regional differences in hereditary breast cancer trends from island to island. In Jamaica, while the mutation rate of \textit{BRCA1/2} was lower than expected, the mutation rate of \textit{PALB2} among unselected patients was the highest reported to date, suggesting that the \textit{BRCA1/2} mutation rate may need to be investigated by further, larger studies, and that \textit{PALB2} is a particularly significant breast cancer susceptibility gene in a Jamaican context. Conversely, in Trinidad and Tobago, \textit{BRCA1/2} mutations were detected at a much higher rate than \textit{PALB2} mutations. However, overall, national screening policies incorporating \textit{BRCA1}, \textit{BRCA2}, and \textit{PALB2} were recommended for both Caribbean regions studied.

A national screening program for both Trinidad and Tobago and Jamaica could begin to be implemented by finding potential carriers from the families of individuals who are diagnosed with breast cancer. If a mutation is detected in one of the three breast cancer susceptibility genes, its pathogenicity can be confirmed by co-segregation analysis. Co-segregation analysis involves calculating the likelihood ratio of a sequence variant to be deleterious. If the variant is found to co-segregate with the disease phenotype in pedigrees, then that makes a strong case for its pathogenicity. This would be the first step in a national screening program that would provide effective early detection measures helpful in reducing late-stand cancers in the Caribbean, thus ameliorating the burden of hereditary breast cancer mortality in the region.
References


36 Lyons TR, Schedin PJ, Borges VF. (2009). Pregnancy and Breast Cancer: when They


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


46 Reeves GK, Pirie K, Beral V et al. (2007). Cancer incidence and mortality in relation
to body mass index in the Million Women Study: cohort study. BMJ 335(7630): 1134.


Lian W, Fu F, Lin Y et al. (2017). The Impact of Young Age for Prognosis by Subtype in Women with Early Breast Cancer. Scientific Reports 7: 10.1038/s41598-017-10414-x.


Hunt KK, Robb GL Strom EA et al. “Primary Prevention of Breast Cancer, Screening


103 Oncotype DX Test [Internet]. BreastCancer.org; c2017 [cited 2017 December 20].
Available from:
http://www.breastcancer.org/symptoms/testing/types/oncotype_dx.

104 EndoPredict Test [Internet]. BreastCancer.org; c2017 [cited 2017 December 20].
Available from:


106 Mammostrat Test [Internet]. BreastCancer.org; c2017 [cited 2017 December 20].
Available from:
http://www.breastcancer.org/symptoms/testing/types/mammostrat.


Broca P. (1866). Traité des tumeurs.


Common Risk Allele Results [Internet]. Clinical Sequencing Exploratory Research; c2017 [cited 2017 December 31]. Available from:


218 Powell SN, Kachnic LA. (2003). Roles of BRCA1 and BRCA2 in homologous


230 Dong Y, Hakimi MA, Chen X et al. (2003). Regulation of BRCC, a holoenzyme complex containing BRCA1 and BRCA2, by a signalosome-like subunit and its


244 Zhang F, Ma J, Wu J et al. (2009). PALB2 Links BRCA1 and BRCA2 in the DNA-


Breast cancer 1, early onset [Internet]. Human Prostate Gene DataBase; c2017


325 Fanconi Anemia, Complementation Group D1; FANCD1 [Internet]. OMIM; c2017 [cited 2017 September 7]. Available from: https://www.omim.org/entry/605724.


332 Zhong Q, Chen CF, Li S et al. (1999). Association of BRCA1 with the hRad50 hMre11-p95 complex and the DNA damage response. Science 285: 747-750.


337  *BRCA1* Gene (Protein Coding) [Internet]. GeneCards; c2017 [cited 2017 September 6]. Available from: http://www.genecards.org/cgi/bin/carddisp.pl?gene=BRCA1.


343  *BRCA1* Gene (Protein Coding) [Internet]. GeneCards; c2017 [cited 2017 September 6]. Available from: http://www.genecards.org/cgi-bin/carddisp.pl?gene=BRCA1.


24.


Minami A, Nakanishi A, Ogura Y et al. (2014). Connection between Tumor Suppressor BRCA1 and PTEN in Damaged DNA Repair. Front Oncol. 4: 318.


Saal LH, Gruvberger-Saal SK, Persson C et al. (2008). Recurrent gross mutations of


431 CDH1 Gene (Protein Coding) [Internet]. GeneCards; c2017 [cited 2017 September 8]. Available from: http://www.genecards.org/cgibin/carddisp.pl?gene=CDH1&keywords=CDH1


ATM Gene (Protein Coding) [Internet]. GeneCards; c2017 [cited 2017 September 6]. Available from: http://www.genecards.org/cgi-bin/carddisp.pl?gene=ATM


PALB2 Gene (Protein Coding) [Internet]. GeneCards; c2017 [cited 2017 September 6]. Available from: http://www.genecards.org/cgi-bin/carddisp.pl?gene=PALB2


XRCC2 Gene (Protein Coding) [Internet]. GeneCards; c2017 [cited 2017 September 6]. Available from: http://www.genecards.org/cgi-bin/carddisp.pl?gene=XRCC2


BARD1 Gene (Protein Coding) [Internet]. GeneCards; c2017 [cited 2017 September 6]. Available from: http://www.genecards.org/cgi-bin/carddisp.pl?gene=BARD1&keywords=BARD1


NBS1 Gene (Protein Coding) [Internet]. GeneCards; c2017 [cited 2017 September 6]. Available from: http://www.genecards.org/cgi-bin/carddisp.pl?gene=NBN

RAD50 Gene (Protein Coding) [Internet]. GeneCards; c2017 [cited 2017 September 6]. Available from: http://www.genecards.org/cgi-bin/carddisp.pl?gene=RAD50&keywords=RAD50


are associated with breast cancer susceptibility. Nature Genetics 47: 643-646.


Seki T, Tada S, Katada T et al. (1997). Cloning of a cDNA Encoding a Novel Importin-α Homologue, Qip1: Discrimination of Qip1 and Rch1 from hSrp1 by


Cancer second leading cause of death in the Caribbean-CARPHA [Internet]. Jamaica Observer; c2015 [cited 2017 October 15]. Available from:


545 Jamaica Population Fact Sheet [Internet]. World Health Organization International Agency for Research on Cancer: GLOBOCAN 2012: Estimated Cancer
In 2012, France. Available from: http://globocan.iarc.fr/Pages/fact_sheets_population.aspx


