1-1-2018

Missing Heritability And Novel Germline Risk Loci In Hereditary Ovarian Cancer: Insights From Whole Exome Sequencing And Functional Analyses

Jaime Lyn Stafford
Wayne State University,

MISSING HERITABILITY AND NOVEL GERMLINE RISK LOCI IN HEREDITARY OVARIAN CANCER: INSIGHTS FROM WHOLE EXOME SEQUENCING AND FUNCTIONAL ANALYSES

by

JAIME L. STAFFORD

DISSERTATION

Submitted to the Graduate School

of Wayne State University

Detroit, Michigan

in partial fulfilment of the requirements

for the degree of

DOCTOR OF PHILOSOHY

2018

MAJOR: MOLECULAR GENETICS AND GENOMICS

Approved By:

_____

Advisor Date

_____

_____

_____

DEDICATION

I would like to dedicate this manuscript to my parents, and my husband who have never shown anything but love and support.

ACKNOWLEGEMENTS

I simply cannot express with words my sheer gratitude towards Dr. Michael A. Tainsky for allowing me to pursue my Doctoral degree under his mentorship. His guidance and trust create such a positive working environment for all involved. I am grateful to all members of the Tainsky lab with a special thanks to Nancy Levin who led patient ascertainment for the registry, was a tremendous help with Sanger confirmation, and always gave such positive energy and kind words. In addition, I would like to thank Sophia Chaudhry who has also helped to ensure the success of this project. Furthermore, I would like to thank my committee members who sacrificed their valuable time to provide me with guidance and insight. I would also like to extend my warmest thanks to Guilherme S. Lopes for his support during the statistical analysis and Viswateja Nelakuditi for his invaluable guidance during the NGS data curation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Ab	Antibody
ATM	Ataxia telangiectasia mutated
BRCA1	Breast Cancer
BRCA2	Breast Cancer
CHK1	Checkpoint Kinase one
CHK2	Checkpoint kinase two
HMMR	Hyaluronan Mediated Motility Receptor
TP53I3	TP53 inducible protein 3
FANCM	Fanconi anemia complementation group M
RAD51D	RAD51 Paralog D
bp	Base pair
cDNA	Complementary DNA
C	Celsius
DNA	Deoxyribonucleic acid
DSB	Double-strand DNA break
EDTA	Ethylenediaminetetraacetic acid
g	Gram
HBOC	Hereditary breast and ovarian cancer syndrome
HDR	Homology directed repair
HRR	Homologous recombination repair
LOH	Loss of heterozygosity
LOF	Loss of function

OVCA	Ovarian cancer
µl	Microlitre
mA	Milliamps
MAF	Minor allele frequency
mM	Millimolar
ml	Millilitre
NHEJ	Non-homologous end joining
NGS	Next generation sequencing
PCR	Polymerase chain reaction
RT	Room temperature
SNP	Single nucleotide polymorphism
siRNA	Silent inhibitory ribonucleic acid
ssDNA	Single stranded DNA
WES	Whole exome sequencing
WGS	Whole Genome Sequencing
WT	Wild-type
VUS	Variant of unknown clinical significance

CHAPTER 1: INTRODUCTION

1.1 Hereditary cancer predisposition

Cancer is a complex disease defined by uncontrolled cell growth that eventually leads to cells invading other nearby organs. It remains one of the leading causes of morbidity and mortality worldwide and places an enormous financial burden on those affected. Startlingly, the estimated financial annual cost of cancer is projected to reach $158 billion by 2020 in the US alone (Mariotto et al., 2011). There are over 100 types of cancer, classified by the organ of origin. What causes this uncontrolled growth is not always clear. The advent of molecular profiling has helped to uncover mechanisms underlying tumor development and at the same time has helped to further classify tumors by sub-type. Decades of research have shown various environmental carcinogens to play a role as well as genetic and epigenetic contributions.

Tumorigenesis is a multistep genetic process and usually begins with a somatic mutation in a tumor suppressor gene (involved in DNA repair, apoptosis, etc.) or oncogene. For tumor suppressor genes, an additional somatic mutation causing a loss of heterozygosity (LOH) of the wild type (WT) allele is often necessary for tumor progression. This observation confirmed Knudson's two-hit hypothesis (Knudson, 1971). Individuals who carry a germline variant in a tumor suppressor gene are at significantly higher risk than the general population as they are born already with the "first hit". These individuals are said to have an inherited cancer risk predisposition (Figure 1).

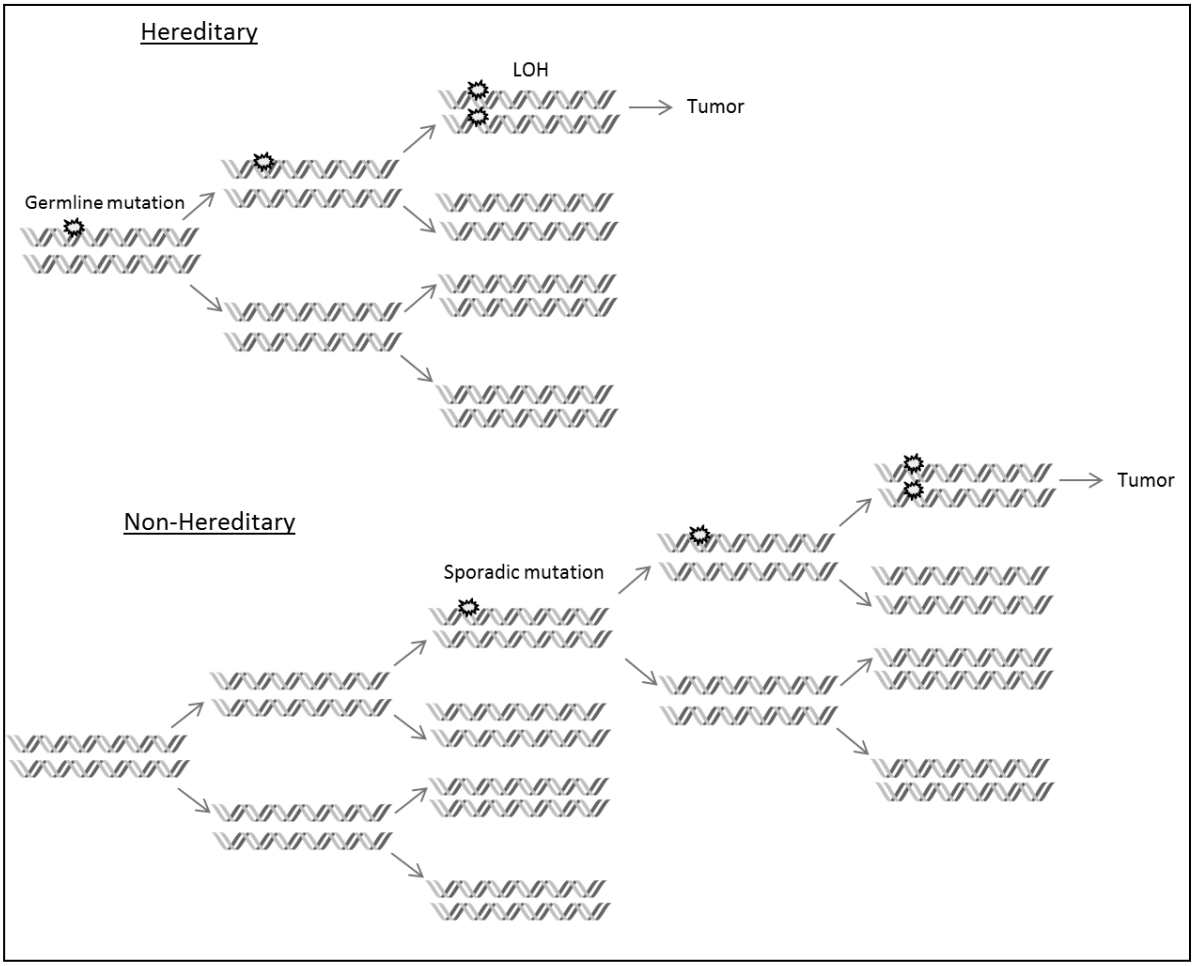


Figure 1. Non-hereditary and hereditary initiation of tumorigenesis by two-hit hypotheses

1.2 Ovarian cancer

Ovarian cancer (OVCA) begins in the ovaries or Fallopian tubes, which are part of the female reproductive system (Figure 2). Ovaries are almond in size and shape and are responsible for ova and female hormone production. Ovarian cancer is less common than other cancers, such as breast or colon cancer, and accounts for approximately 3% of all of cancers in women (National Cancer Institute, 2013). Still, it is estimated that over 21,000 new cases of ovarian cancer occur yearly, in the United States alone (American Cancer Society, 2014). Yet, despite being ranked as the 11th most common cancer in women, it is the 5th leading cause of cancer-related death among women, and is the top cause of female reproductive cancer deaths (Bell et al., 2011). The five year survival rate at diagnosis averages just 44.2% (American Cancer Society, 2014). The high mortality rate is mostly due to poor early detection as only 15% of patients are diagnosed in early, more treatable stages (National Cancer Institute, 2013). Accurate risk prediction holds promise as prophylactic measures can be taken before the cancer develops (Bast, Hennessy, and Mills, 2009).

Recent studies suggest that at least 25% of epithelial ovarian cancer cases arise due to an inherited risk factor (Walsh et al., 2011). Hereditary breast and ovarian cancer (HBOC) syndromes are, for the most part, autosomal dominant genetic disorders in which germline mutations elevate lifetime risk of developing breast or ovarian cancer up to as much as 87% and 49% respectively (Antoniou et al. , 2003; Risch et al. , 2001; Claus et al. , 1996). The risk among the general population is 12% for breast and 1.4% for ovarian cancer (Plevová et al. 2009). Therefore, women with a personal or family history of OVCA and/or young onset and/or multiple cases of breast cancer are counseled to consider genetic screening per guidelines of the National Comprehensive Cancer Network (NCCN) (Genetic/Familial High-Risk Assessment:

Breast and Ovarian www.nccn.org). Current testing panels mostly feature genes involved in DNA repair and cell cycle control, such as *BRCA1* and *BRCA2*, which explain the majority of inherited ovarian and breast cancer, as well as 22 other genes including *ATM, BARD1, BRCA1, BRCA2, BRIP1, CDH1, CHK2, MLH1, MRE11A, MSH2, MSH6, MUTYH, NBN, PALB2, PMS2, RAD50, RAD51C, RAD51D, SMARCA4,* and *TP53*. Additional genes tested that are not directly involved with DNA repair or cell cycle control include; *EPCAM, NF1, PTEN* and *STK11*.

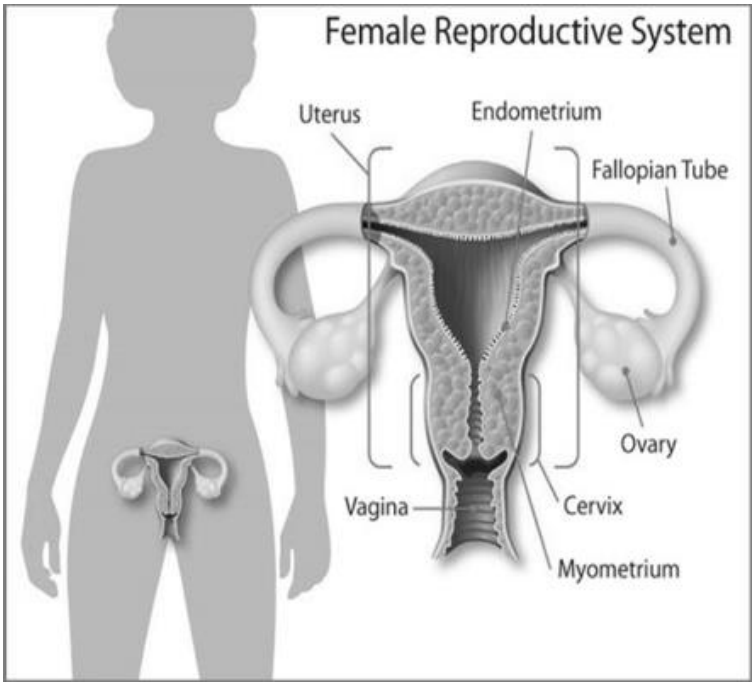


Figure 2. Diagram of the female reproductive system

Source: *https://www.womenshealth.gov/cancer/ovarian-cancer*. This image is free of copyright restrictions and may be copied, reproduced, or duplicated without permission of the Office on Women's Health in the U.S. Department of Health and Human Services

1.2.1 Diagnostics and detection

While significant advances in diagnosis and treatment for various other forms of cancer have led to excitement in the field of oncology, OVCA mortality rates have only slightly decreased since "the war on cancer" was officially declared in 1971 (NIH Surveillance, Epidemiology and End Results Program: Cancer of the Ovary - Cancer Stat Facts, 2017). This is mostly due to a lack of advancement in early detection since survival rates depend greatly on the stage at which the cancer is diagnosed. Women diagnosed before the cancer has had a chance to metastasize have a dramatically increased five-year survival rate relative to those diagnosed at a more advanced stage. The 5-year survival rate is over 90% with diagnosis at stage I and standard treatment. Unfortunately, less than 20% percent of OVCA patients are diagnosed in the early stages of the disease (National Cancer Institute, 2013). Ovarian cancer is difficult to detect early since a woman may be asymptomatic until the cancer has advanced to Stages 3 and 4. Unlike breast cancer, tumors that grow in the ovaries or Fallopian tubes are not readily detectable by self-examination. When symptoms do appear, they are often vague and may not illicit immediate medical care. Symptoms include: abdominal pain, vaginal bleeding, weight loss, abnormal periods, back pain, bloating, nausea and vomiting, etc (NIH, 2017). Therefore, accurate risk assessment to identify likely cases prior to occurrence holds much promise since prophylactic surgery can virtually eliminate a woman's chances of developing OVCA.

1.2.2 Risk factors

Today, in the United States, the average woman's lifetime risk of developing invasive ovarian cancer is 1 in 75 (National Cancer Institute 2013; Risk Factors - Ovarian Cancer Research Fund Alliance, 2017). While women of all ages are susceptible to developing ovarian cancer, incidences are highest in women 55-64 years of age (National Cancer Institute, 2013;

Risk Factors - Ovarian Cancer Research Fund Alliance, 2017). There are, in addition to age, multiple factors that affect a women's lifetime risk. Caucasian ancestry, nulliparity, infertility, the use of hormonal replacement therapy, and obesity have all been found to correlate with an increased risk (National Cancer Institute, 2013; Risk Factors - Ovarian Cancer Research Fund Alliance, 2017). The reproductive history is thought to be important to ovarian cancer risk since the risk increases with the amount of menstrual cycles a woman has during her fertile years. Therefore, a woman is at an increased risk if she has never had any children, has begun menstruation before the age of 12, started menopause after 50, or has never used oral contraceptives or undergone tubal ligation (National Cancer Institute, 2013; Risk Factors - Ovarian Cancer Research Fund Alliance, 2017).

However, none of these factors are strong enough to predict the risk of ovarian cancer occurrence with enough conviction as to prompt prophylactic surgery. The single most influential factor on a woman's risk is her family history (National Cancer Institute, 2013; Risk Factors - Ovarian Cancer Research Fund Alliance, 2017; Chun and Ford, 2012; Trifonov, Todorova, and Uzunova, 2001). While the average woman’s lifetime risk is just 1.4%, those with a first degree relative diagnosed with OVCA have a 5% lifetime risk (Centers for Disease Control and Prevention, 2017). Therefore, women with a family history of ovarian cancer are encouraged to undergo genetic risk assessment.

1.3 Genetic risk assessment

Hereditary breast and ovarian cancer (HBOC) syndromes are for the most part autosomal dominant genetic disorders where germline mutations elevate lifetime risk of developing breast or ovarian cancer up to 87% and 49% respectively (Risch et al., 2001; Claus et al., 1996). Genetic testing can identify those at risk before the cancer develops, which in some cases leads

to prophylactic measures. For those already with a diagnosis, testing for germline mutations can inform risk of additional cancers as well as identify risk in unaffected relatives who may also carry the same mutation. Therefore, identifying those at increased risk due to genetic inheritance can lead to improved clinical outcomes (Narod et al., 2013). A diagnosis of HBOC is considered when either ovarian or breast cancer occurs on the same side of a family for multiple generations or multiple first-degree relatives (Kobayashi et al., 2013). Specifically, the likelihood of familial risk for HBOC increases with each of the following criteria (Clinical, Guidelines, and Guidelines 2018):

- A diagnosis of breast or OVCA under the age of 45
- OVCA at any age
- A blood relative diagnosed with breast cancer before 50 years of age
- Breast/OVCA across multiple generations on the same side of the family
- A second diagnosis of breast cancer or ovarian cancer in the same individual
- Breast cancer in a male blood relative
- Breast or ovarian and pancreatic cancer all on the same side of the family
- A history of cancer in a family of Ashkenazi Jewish ancestry

Two genes commonly associated with HBOC are *BRCA1* and *BRCA2*, while other less common mutations have been found in tumor suppressor genes such as *TP53*, *PTEN*, *CDH1*, *ATM*, *CHK2* or *PALB2*, etc. (Kobayashi et al., 2013). Another inherited syndrome, Lynch Syndrome (also known as hereditary nonpolyposis colorectal cancer or HNPCC) has also been linked to ovarian cancer with a 12% lifetime risk (National Comprehensive Cancer Network, 2016). Because of the heterogeneity of inherited OVCA risk, testing clinics currently offer large genetic sequencing panels which include all 24 genes known to associate with increased risk of

breast, ovarian, and/or uterine cancers; *ATM, BARD1, BRCA1, BRCA2, BRIP1, CDH1, CHK2, EPCAM, MLH1, MRE11A, MSH2, MSH6, MUTYH, NBN, NF1, PALB2, PMS2, PTEN, RAD50, RAD51C, RAD51D, SMARCA4, STK11, TP53*.

1.4 Genetic Penetrance

The proportion of individuals carrying a variant (allele) of a gene that also express the trait associated with that variant is referred to as penetrance. The degree of penetrance for disease associated genetic loci varies greatly. Some mutations have complete penetrance, meaning all individuals with the mutation have or will develop the disease (e.g Cystic Fibrosis, Huntington's disease). Most cancer-associated risk loci have incomplete penetrance, and many carriers of the risk mutation may never develop cancer. *BRCA1* is the most highly penetrant gene in hereditary breast and ovarian cancer syndrome with an 80% lifetime risk of breast cancer and up to a 50% lifetime risk of OVCA (Risch et al., 2001; Claus et al., 1996; Hereditary Breast and Ovarian Cancer | Cancer.Net, 2017). Because clinical decisions such as prophylactic surgery and chemo-preventative therapies are made based on risk assessment, accurate penetrance estimates of cancer-associated loci is invaluable. Currently, genes known to be mutated in hereditary cancer syndromes are referred to as having high, moderate, or low penetrance (Table 1, Stanislaw C., 2016). Unfortunately, penetrance is particularly difficult to determine for many cancer-associated risk loci as pathogenic variants in these highly conserved genes tend to be rare. In addition, environmental factors as well as other genetic and epigenetic alterations can contribute, meaning penetrance can vary based on the carrier. Therefore, germline mutations in cancer-associated genes are often described as having an unknown penetrance.

PENETRANCE	GENE	CANCER RISK and GUIDELINES FOR CLINICAL MANAGEMENT
High	*BRCA1* *BRCA2* *APC* *PTEN* *TP53* *MLH1* *MSH2* *MSH2* *PMS2* *STK11* *CDH1* *MUTHY*	Strong evidence for increased risk with well-defined risk profiles, prophylactic surgery or other preventative measures advised
Moderate	*ATM* *CHEK2* *PALB2*	Moderate evidence for increased risk of certain cancers, increased surveillance advised
Low/unknown	*RAD50* *RAD51C* *RAD51D* *BRIP1* *BARD1* *POLE* *POLD1*	Varied evidence with clinical management based on personal and family history

Table 1. Penetrance varies by gene and is often unknown. Table adapted from "Genetic evaluation and testing for hereditary forms of cancer in the era of next-generation sequencing" Stanislaw C., 2016.

1.5 *BRCA1* and *BRCA2*

Both *BRCA1* (breast cancer type one susceptibility) and *BRCA2* (breast cancer type two susceptibility) are tumor suppressor genes necessary for genomic stability. Carriers of a pathogenic mutation in either gene are at a greatly increased risk of multiple types of cancer, most notably breast and ovarian. These two genes account for 5 to 10% of all OVCA cases (Ramus and Gayther, 2009). In the general population, approximately 1 in 400 people have a *BRCA1/2* mutation that leads in increased cancer risk (PDQ Cancer Genetics Editorial Board 2002). The prevalence of *BRCA1/2* pathogenic mutations is particularly high for certain ethnicities, specifically among Ashkenazi Jews where 1 in 40 are carriers of pathogenic founder mutations (Robles-Díaz et al., 2004). Both genes have been well characterized and are involved in numerous cellular processes important for genomic stability.

BRCA1 encodes for an E3 ubiquitin-protein ligase that facilitates a diverse range of cellular processes including DNA damage repair, cell cycle control, apoptosis, transcriptional regulation and embryonic development (W. Wu et al., 2008). BRCA1 interacts with numerous proteins by taking part of large complexes required for these pathways. For instance, BRCA1 has been shown to interact with RNA polymerase II for regulation of p21 in response to DNA damage (Moisan and Gaudreau, 2006), is involved in the MRE11-RAD50-NBS1 (MRN) complex (L. Chen et al., 2008), as well as binds BRCA2 and RAD51 for homology directed repair of double stranded DNA breaks (J. J. Chen et al., 1999). BRCA1 has also been shown to regulate chromatin remodeling via the SWI/SNF complex (Bochar et al., 2000) and ubiquitinates RBBP8 for CHK1 mediated G2/M cell cycle control (Yarden et al., 2012).

Similarly, BRCA2 is also involved in numerous cellular processes required for genomic stability. These biological processes include DNA repair, cell division, histone acetylation, and

replication fork maintenance. BRCA2 selectively binds to ssDNA and mediates HRR by aiding in RAD51 ssDNA assembly and stabilization. This enables RAD51 to displace replication protein-A (RPA) which binds single strand DNA (ssDNA) during the initial steps of DNA repair (J. J. Chen et al., 1999). BRCA2 forms a complex with PALB2 to direct the localization of POLH to collapsed replication forks (Buisson et al. 2014). In addition, this interaction is part of a complex including RAD51C and involved in DNA repair by HRR (Sy et al., 2009). BRCA2 also regulates centrosome duplication via NPM1 (H. F. Wang et al., 2011). The BRCA2 protein likely has additional functions yet to be verified. For example, it is suspected to play a role in S phase checkpoint activation (Yoshida and Miki, 2004) and prevent R-loop DNA damage incurred through the transcription process (Bhatia et al., 2014).

1.6 Missing heritability

For complex traits such as susceptibility to cancer it is possible to establish an estimate of how much of the phenotypic variance is due to genetic inheritance. One classic means of determining the heritability of a trait is through twin studies that measure the similarity of a trait between monozygotic twins as compared to dizygotic twins. Monozygotic twins are derived from a single fertilized egg and thus share the same genetic material while dizygotic twins, formed from separate eggs, share about half of their genes. Because both sets of twins share the same environment, at least in childhood, this approach controls for much of the environmental contribution to cancer risk. If monozygotic twins resemble each other more than dizygotic twins for a particular trait, then the heritability of that trait is estimated as twice the observed difference.

Using this approach, a high amount of heritability has been observed in various types of cancers, including melanoma (58%), prostate (57%), ovary (39%), and breast (31%) amongst

others (Mucci et al., 2016). However, as with many multifactorial diseases, there is a gap in our knowledge of the genetics underlying the predisposition. Often, the known genetic variation attributed to a disease do not account for its total estimated heritability. This is referred to as *'the missing heritability issue'*. For certain diseases, this unexplained heritability often limits how informative genetic testing can be. Specifically, while ovarian cancer is demonstrated to have a strong genetic component, the known risk loci cannot account for the majority of the familial risk. Therefore, many women with compelling personal and or family histories regularly test negative for the currently described susceptibility loci (Stafford et al., 2017).

1.7 Variants of unknown significance

Rare and private mutations are likely to account for much of the missing heritability in ovarian cancer. This hypothesis is supported by the fact that many women who undergo genetic testing are found to have a "variant of unknown significance" (VUS) (Stafford et al., 2017; Towler et al., 2013; Domchek and Weber, 2008). The term "*VUS*" is used to describe a rarely seen mutation, unannotated in its functional consequence on the protein or disease risk. Assessing these variants is crucial for accurate risk assessment, prevention and targeted therapy. Yet we currently lack a means to functionally validate private mutations in a clinical setting (Towler et al., 2013). The American College of Medical Genetics and Genomics (ACMG) provides detailed guidance as to the interpretation of genomic variants (Richards et al., 2015) (Table 2). However, in the case of a missense change, it is difficult to obtain strong evidence for pathogenicity, especially when rare. Therefore, the vast majority of compelling sequence changes remains annotated as having only moderate evidence as to its pathogenicity.

ACMG Evidence for Pathogenicity:	
VERY STRONG	• Loss of function variant in a cancer-associated gene
STRONG	• Functional assay supportive of damaging effect • Increased prevalence in cases versus controls (O.R. >5) • Co-segregation with disease in multiple affected family members
MODERATE	• In a mutational hot spot or well-established functional domain • Extremely low frequency or absent from ExAC, 1000 genomes • In-frame deletions/insertions • Co-segregation with disease in multiple affected family members • Multiple lines of computational evidence

Table 2. ACMG Evidence for Pathogenicity. Simplified and adapted from "Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology"

1.8 Implicating novel risk loci

When a trait or disease is suspected to be inherited, there are various approaches that can be employed to identify the genes and variants responsible. One approach is through linkage analysis, a hypothesis-neutral means to find segments of DNA which all affected persons of a family share but the healthy relatives do not have. This is possible because genetic variants that are close in proximity tend to be inherited together. The analysis begins with genotyping genetic "markers" (highly variable regions of the DNA) on each chromosome for both affected and unaffected individuals of a family. Once a region is identified as shared between and unique to those affected, additional markers are added to further narrow the search within this region until a specific gene and genetic variant is implicated. This method is most appropriate when attempting to map monogenic disease mutations and requires large families.

For complex traits and disease predisposition, a better approach to implicating genetic variants is through association, such as a *Genome Wide Association Study* (GWAS). These studies look for genomic variants that are statistically more prevalent in cases versus controls by genotyping across the entire genomes of hundreds or thousands of individuals both with and without the trait/disease of interest. GWA studies rely heavily on the "common disease, common variant" (CDCV) assumption and are typically designed to exclude SNPs with a MAF < 5% (Visscher et al., 2012). They lack the statistical power necessary to detect rare variants, which are more likely to have a high effect. Therefore, GWA studies are best for implicating common and low risk alleles that are not clinically actionable.

High-throughput sequencing using next generation technology allows for a more efficient and unbiased approach in the discovery of novel cancer predisposition loci and has helped to determine the frequency of germline mutations in HBOC. However, because of the rarity of

ovarian cancer, large cohorts are obtained by recruiting participants that are not selected based on family history. This means that, in large studies, most of the underlying etiology is sporadic, and the majority of the causal variants uncovered are in *BRCA1* and *BRCA2*, genes with already well-established roles in OVCA. For example, a recent publication reports the results of sequencing 1915 patients with OVCA and found that of these cases, 347 (18%) carried pathogenic germline mutations in genes associated with OC risk, 80% of which were in *BRCA1* and *BRCA2* (Norquist et al., 2015). Because patients were not selected for age or family history, the vast majority of cases in this cohort are sporadic and not due to an underlying inherited risk. This type of study, while impressive in size, is limited to estimating the prevalence of known pathogenic mutations in the study population.

Hereditary OVCA is a rare disease with a high degree of genetic heterogeneity. Therefore, despite the fact that any women with a diagnosis of OVCA is considered at risk of possible germline risk, the best study design for identifying novel putative risk loci includes a much stricter selection of patients. The most appropriate study sample would include only OVCA patients with a suspicious family history indicating genetic inheritance of risk and would exclude those with known pathogenic variants (Stafford et al., 2017). Candidate risk loci identified by the whole exome/genome sequencing of this sample would then be assessed by both bioinformatics and functional analysis.

1.9 Gene panel vs whole exome sequencing for diagnostic purposes

Because of the immense heterogeneity of inherited cancer risk, current testing is performed by analyzing a panel of associated risk genes to increase the chance of finding a causal genetic variant. However, despite the introduction of larger and more inclusive gene panels, issues of low diagnostic yield remain, and inconclusive results have brought an additional

burden of stress on clinicians, patients, and relatives. Our knowledge regarding inherited cancer risk is still very incomplete. Therefore, testing panels are frequently being updated and subsequently outdated and many who question the utility of panels favor a more practical whole genome/exome approach by which genes can be analyzed *post hoc* without the need for additional testing. The future of genetic testing seems poised for this approach, but hesitation remains for various reasons.

One valid argument against employing exome sequencing in place of panel testing is that it would sacrifice quality for quantity. High quality variant calls are necessary for accurate clinical genetic diagnosis and whole exome sequencing (WES) is not ideal for detection of variants in regions high in GC content, with many repeats, or with homology to other areas of the genome (Sekhar et al., 2014). In contrast, panel testing focuses on a small number of genes allowing for the luxury of more coverage, greater read depth, and thus higher quality variant calls. Therefore, panels are traditionally believed to have superior detection of pathogenic variants and thus a better diagnostic yield. However, as our technology advances and bioinformatic pipelines for variant calling improves, the quality of WES data will inevitably reach comparable clinical sensitivity if it has not yet already. A recent study assessing the coverage in 100 samples demonstrated that 99.7% of pathogenic variants were detectable by WES at clinical sensitivity, and all had at least some coverage on exome sequencing (Laduca et al., 2017). Another study compared the diagnostic performance of WES to two panels, TruSight Cancer (94 genes) and another custom panel of 122 genes and identified a similar number of variant calls (amongst shared genes) despite greater average read depth in panels (Feliubadaló et al., 2017).

Another important consideration besides call quality is cost. However, the cost discrepancy between WES and panels is narrowing and the long-term benefits of exome sequencing may, in some cases, outweigh the difference. While sequencing exomes is more expensive than sequencing panels, they provide additional patient data regarding genes that may one day be clinically relevant and is accessible without the need for retesting. Panels can quickly become outdated with each novel gene discovery and insurance will often only cover genetic testing once.

Yet another valid concern for molecular diagnostic labs is testing turn-around time. Many smaller labs may not be set up for WES on so many samples and sometimes immediate clinical decisions are based on mutational status. Therefore, when to choose exome sequencing over panel testing should be made on a case by case basis. Currently, as a first step into this new era, WES may only be necessary in cases suspected to have inherited risk. For instance, NCCN guidelines stipulate that women diagnosed with ovarian cancer are eligible for genetic testing regardless of family history ("National Comprehensive Cancer Network. Panel Members Genetic/Familial High-Risk Assessment: Breast and Ovarian (Version 1), 2018). While panel testing is likely sufficient for most cases, those patients with a family history of breast or ovarian cancer may benefit more from whole exome sequencing.

1.10 DNA repair and genomic integrity

The integrity of an organism's genetic material is essential to its survival. However, our DNA is constantly being damaged during the process of cellular metabolism as well as by external environmental assaults such as radiation. Consequently, all forms of life have developed the ability to identify and repair these genomic lesions. While there exist many types of DNA lesions, the most harmful is the double strand break (DSB). DSBs occur due to

exposure to ionizing radiation (IR), various chemotherapeutic reagents, reactive oxygen species (ROS) generated by cellular metabolic processes and in certain circumstances during DNA replication (Mehta and Haber, 2014). DSBs that go unrepaired can lead to a deletion, amplification or even translocation of genetic material, potentially at a locus encoding for a tumor suppressor or oncogene. Therefore, effective repair of DSBs by either homologous recombination repair (HRR) or non-homologous end joining (NHEJ) is crucial to avoid tumorigenesis.

1.10.1 Homologous recombination repair

Homologous recombination repair (HRR) is a highly conserved and non-error prone process that corrects DSBs using the sister chromatid as a template for gene conversion (Figure 3). Therefore, this mechanism can only be employed during G2 and S phase of the cell cycle, when a homologue is available. Shortly after the double strand lesion occurs, HRR is initiated by the phosphorylation of H2AX at Ser139 by ATM or ATR (Helt et al., 2005; Rogakou et al., 1998). Phospho-H2AX (γH2AX) recruits various proteins to the damaged DNA, most notably MRE11, RAD50 and NBS1 which form the MRN complex. This complex binds the broken ends of the DNA for stability, likely preventing additional breakage in addition to initiating DNA strand resection by endonuclease MRE11 (Lamarche, Orazio, and Weitzman., 2010). The resection creates single-stranded DNA (ssDNA) which is stabilized by Replication Protein A (RPA) and acts as a probe in the search for a homologous sequence. Next, the broken strand invades the sister chromatid through the action of RAD51, which replaces the RPA through the action of BRCA2 (Buisson et al.. 2010). RAD51 then replaces RPA and forms filaments on the DNA necessary for the strand exchange between homologous sequences (Buisson et al. 2010). Synthesis of the missing DNA segment is filled in by replication machinery PCNA and

polymerase δ and ε (Sneeden et al., 2013). Finally, the connected segments of the two DNA molecules (Holliday junction) are resolved by nucleases.

1.10.2 Non-homologous end joining

Another pathway for DSB repair is Non-Homologous End Joining (NHEJ). During this process, breaks in DNA are corrected by a direct ligation of the blunt ends and thus do not require a homologue to serve as a template. While this mechanism is more error prone, it is the most reliable mechanism for DSB repair in post-mitotic cells. NHEJ involves three basic steps: DNA DSB damage recognition, processing of break ends, and strand ligation (Figure 3). In NHEJ, DSBs are quickly recognized by heterodimer KU70/KU86 (KU80) which act to both protect the strand ends from degradation as well as recruit DNA-PKcs, the catalytic subunit of DNA dependent protein kinase (Jin and Weaver, 1997). End processing may occur before ligation and may involve several enzymes including a nuclease (ARTEMIS), polymerases, polynucleotide kinase/phosphatase (PNKP), and Aprataxin (APTX) among others (Povirk, 2012). For ligation of the processed break ends, both DNA ligase IV and XRCC4 are recruited and activated via phosphorylation by DNA-PKcs (Povirk, 2012).

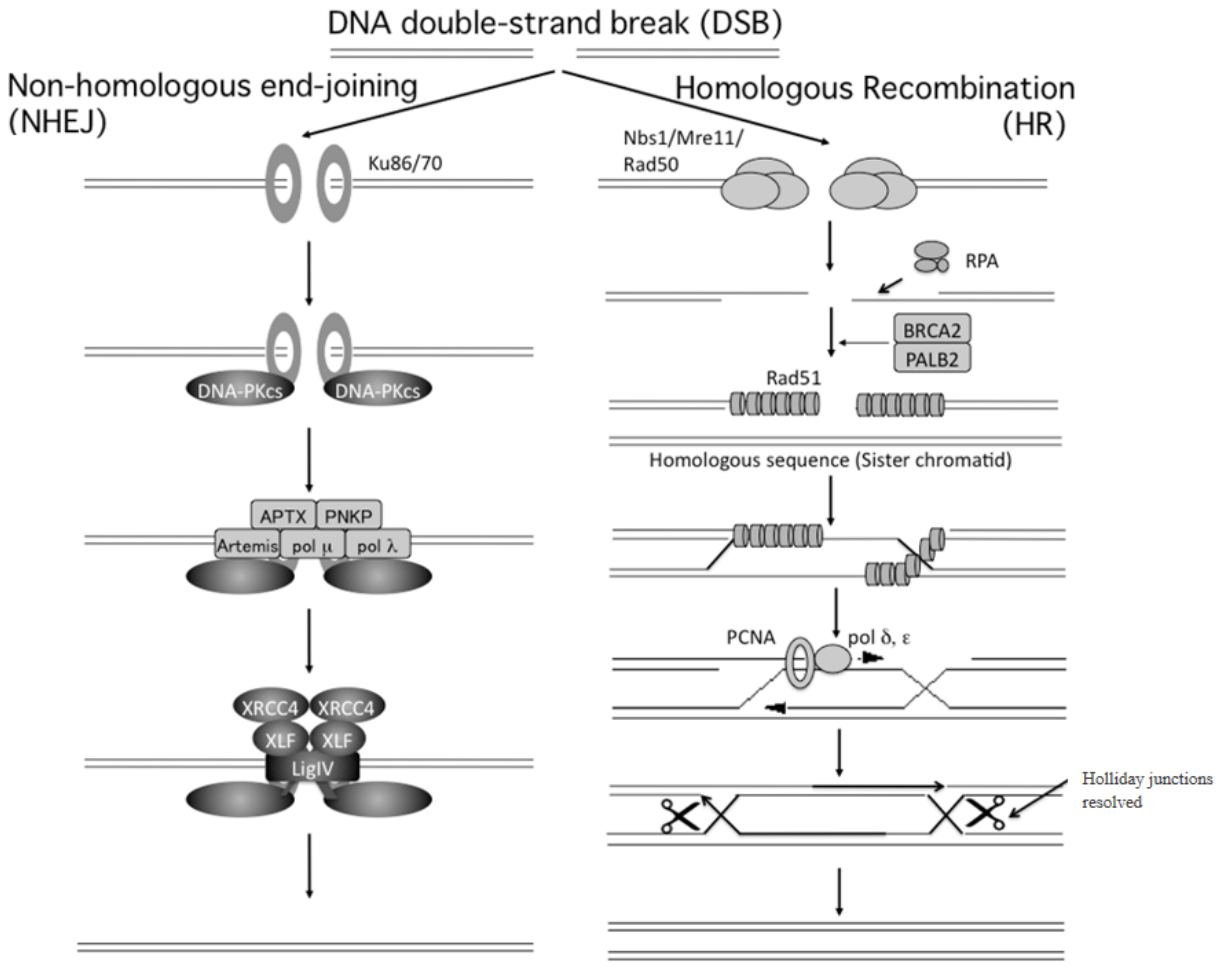


Figure 3. Schematic illustration of the repair of a DNA DSB by homologous recombination repair (HRR) and non-homologous end joining (NHEJ). Image Source: *https://www.intechopen.com/books/new-research-directions-in-dna-repair/radiosensitization-strategies-through-modification-of-dna-double-strand-break-repair*. Permission of image use granted by Dr. Yoshihisa Matsumoto, Tokyo Institute of Technology.

1.11 Targeted therapy in clinical cancer care

Our greater understanding of the molecular genetics underlying tumorigenesis has led to a whole new field of precision medicine in which therapies are developed to target specific "drivers" or deficiencies that encourage cancer progression. What initiates or "drives" the process of tumorigenesis and metastasis differs by cancer subtype and even by individual tumor. Tumor sequencing has thus become a routine practice in clinical cancer care to inform therapeutic decisions. For instance, the most common driver of tumorigenesis in breast and ovarian cancer is genomic instability due to inefficient DNA repair by homologous recombination (Spring and Perspect 2015; van Gent and Kanaar 2016). Platinum-based agents, the first-line therapy for these tumors, exploit this deficiency by causing DSBs vis intra and interstrand DNA crosslinking (Ph and Andrea, 2016). Cancer cells are most sensitive to this drug as they lack the ability to repair these breaks and damage will accumulate leading to eventual cell death.

Another targeted therapy often used in conjunction with platinum in breast and ovarian cancer is polyADP-ribose polymerase (PARP) inhibitors (PARPi). The PARP family of proteins is involved in a form of ssDNA repair known as Base Excision Repair (BER). Cancer cells deficient in other DNA repair mechanisms are more dependent on the PARP proteins. By inhibiting this pathway, the cell loses its ability to repair DNA leading to synthetic lethality. Tumor cells with mutant *BRCA1* and/or *BRCA2* have demonstrated to be up to 1000 times more sensitive to PARPi as compared to WT cells (Farmer et al., 2005; Bryant et al., 2005). In clinical trials, the use of PARPi has shown to improve progression-free survival when added to the treatment of women with breast or ovarian cancer responsive to platinum (Robson et al. 2017; Ledermann, 2016; Evans and Matulonis, 2017). More recently, investigators have identified an

additional mechanism of action for PARPi. PARP proteins were found to localize at DNA damage sites and become trapped on the DNA when bound to an inhibitor, blocking DNA replication, and causing cell toxicity (Lord and Ashworth, 2017).

CHAPTER 2: MATERIALS AND METHODS

2.1 Sample ascertainment and description

Study samples were acquired through the Karmanos Cancer Institute Genetic Registry (KCIGR), an IRB approved biospecimen repository comprising females with a personal or family history of breast and/or ovarian cancer and at elevated risk of harboring a *BRCA1/2* mutation. Over 800 DNA samples from breast and/or ovarian cancer patients were collected spanning the years of 1999-2013, when HBOC genetic screening was limited to *BRCA1/2* and risk assessment was performed using BRACAPRO and Myriad II, which were the standard of care during the duration of accrual. BRCAPRO is a computer-based Bayesian probability model that uses breast and/or ovarian cancer family history to determine the probability that a *BRCA1* or *BRCA2* mutation accounts for the pattern of these cancers in the family (Parmigiani, Berry, and Aguilar, 1998). Key attributes of consideration include the population prevalence of *BRCA* mutations, age-specific penetrance, and Ashkenazi Jewish heritage. Myriad II is a set of prevalence tables categorized by ethnic ancestry (Ashkenazi Jewish or non-Ashkenazi Jewish), breast cancer age of onset (age $\leq$50 years), and the presence of ovarian cancer, in the patient and/or first- or second-degree relatives. Myriad II is based on historical test data from the Myriad Genetic Laboratories clinical testing service (Frank, 1999).

Through the KCIGR biospecimen repository, we obtained 89 DNA samples from high risk Caucasian women with a personal history of OVCA. Participants were either confirmed *BRCA1/2* mutation carriers or *BRCA1/2* negative after full gene sequencing, BART (BRCAnalysis rearrangement test) or testing for the three common Ashkenazi Jewish mutations (Myriad Genetics Laboratories, Salt Lake City, Utah). Participants testing positive for pathogenic *BRCA1/2* germline mutations were excluded from the study sample. The final

sample consisted of 48 *BRCA1/2* mutation negative Caucasian OVCA patients from 47 families (one mother-daughter pair). All subjects provided written informed consent for the collection of blood samples and access to medical records. The protocol (HIC#024199MP2F(5R)) was approved following Full Board Review by the Human Investigation Committee at Wayne State University, Detroit, Michigan.

Information regarding tumor histology, tumor grade, age of diagnosis, and family history of study sample is summarized in Table 3. Tumor histology from study sample patients included serous (n=26), endometrioid carcinoma (n=5), mixed (n=4) adenocarcinoma (n=2), mucinous (n=1), clear cell (n=1), and undefined (n=9). Tumor grade included grade 2 (moderately differentiated, n=6), grade 3 (poorly differentiated, n=24), and grade 1 (well differentiated, n=1). Ovarian cancer was the primary diagnosis for 43 patients, while four had a primary diagnosis of breast cancer and one of cervical cancer followed by a secondary OVCA diagnosis. Of those with a primary OVCA diagnosis, six had a secondary cancer diagnosis: two breast, two colon, one uterine and one melanoma.

Mean Age at Diagnosis	52.8 (yrs.)	25-81 (range)
Histology	**n=**	**%**
Serous	26	54
Endometrioid	5	10
Mixed	4	8
Adenocarcinoma, NOS	2	4
Clear Cell	1	2
Mucinous	1	2
Unknown	9	19
Stage	**n=**	**%**
I	8	17
II	5	10
III	23	48
IV	3	6
Unknown	9	19
Grade	**n=**	**%**
Grade 1- well differentiated	1	2
Grade 2- moderate	6	13
Grade 3- poor	24	50
Unknown	17	35
Personal and Family History	**n=**	**%**
personal BC/OVCA diagnosis < 50 yrs. of age	15	31
personal second primary cancer diagnosis	12	25
personal/family history of BC	31	65
family history of OVCA	14	29
family history of epithelial cancer	47	98

Table 3. Tumor histology, tumor grade, age of diagnosis and family history of study sample

2.2 Whole exome sequencing

DNA from peripheral blood samples was isolated by the Karmanos Applied Genomics Technology Center, Detroit, MI using QIAamp DNA mini kit (Qiagen) and whole exome sequencing (WES) was performed using Nextera Rapid Capture Kit. Samples were demultiplexed using Illumina's CASAVA 1.8.2 software (Kumar et al., 2014). Read quality was assessed with FastQC (*http://www.bioinformatics.babraham.ac.uk/projects/fastqc/2*) and alignment to the human reference genome (hg19) (Lander et al., 2001) was performed using Burrows Wheeler Aligner (BWA) (Li and Durbin, 2009). PCR duplicates were removed using Samtools, (Li et al., 2009) and subsequent local realignment, Qscore recalibration, variant calling and filtering was performed using Genome Analysis Toolkit (GATK) (DePristo et al., 2011) Unified Genotyper and Haplotype Caller. SNPs were filtered out if: 1). Four or more alignments have mapping quality = 0 and the number of alignments that mapped ambiguously were more than 1/10 of all alignments for the given SNP. 2). SNP is represented by less than 5 reads 3). SNP quality is below 50 4. QD score (variant confidence) is below 1.5. Variant files were constructed using Genome Trax BIOBASE biological databases analysis software (http://www.biobase-international.com) and annotated with Illumina BaseSpace VariantStudio application v2.2.4 (www.illumina.com) and snpEff (Cingolani et al., 2012).

2.3 Panel and candidate genes analyzed

Genes analyzed for potential risk variants included those currently featured on HBOC genetic testing panels by Ambry OvaNext and Myriad MyRisk: *BRCA1, BRCA2, MLH1, MSH2, MSH6, PMS2, EPCAM, APC, MUTYH, CDKN2A, CDK4, TP53, PTEN, STK11, CDH1, BMPR1A, SMAD4, PALB2, CHK2, ATM, NBN, BARD1, BRIP1, RAD51C, RAD51D,* in addition to 155 non-panel candidate genes important to DNA damage response or cell cycle regulation

and 64 genes listed as having disease causing mutations associated with OVCA in HGMD. A full list of non- panel candidate genes analyzed is provided in Table 4.

DNA repair cell cycle control genes:

APEX1	*MGMT*	*RAD50*	*XRCC2*	*CCND2*	*CHK1*	*MKI67*
APEX2	*MLH1*	*RAD51*	*XRCC3*	*CCNE1*	*CKS1B*	*MNAT1*
BLM	*MLH3*	*RAD51C*	*XRCC4*	*CCNF*	*CKS2*	*RAD9A*
DDB1	*MPG*	*RAD52*	*XRCC5*	*CCNG1*	*CUL1*	*RB1*
DDB2	*MSH2*	*RECQL*	*XRCC6*	*CCNG2*	*CUL2*	*RBBP8*
ERCC1	*MSH3*	*RECQL4*	*ZFP276*	*CCNH*	*CUL3*	*RBL1*
ERCC2	*MSH5*	*RECQL5*	*FANCM*	*CCNT1*	*DDX11*	*RBL2*
ERCC3	*NEIL1*	*RFC1*	*HMMR*	*CCNT2*	*DIRAS3*	*SERTAD1*
ERCC4	*NTHL1*	*RFC2*	*POLK*	*CDC16*	*DNM2*	*SKP2*
ERCC5	*OGG1*	*RFC3*	*POLQ*	*CDC2*	*E2F4*	*SKP2.*
ERCC6	*PARP1*	*RFC4*	*REC8*	*CDC20*	*GADD45A*	*SUMO1*
ERCC8	*PCNA*	*RFC5*	*ABL1*	*CDC34*	*GTF2H1*	*TFDP1*
EXO1	*PMS1*	*RPA1*	*ANAPC2*	*CDK2*	*GTSE1*	*TFDP2*
FANCA	*POLB*	*RPA2*	*ANAPC4*	*CDK5R1*	*HERC5*	*UBA1*
FANCC	*POLD1*	*RPA3*	*ATR*	*CDK5RAP1*	*HUS1*	*TP53I3*
FANCE	*POLD2*	*SMUG1*	*BAX*	*CDK6*	*KNTC1*	*TP53BP1*
FANCF	*POLD3*	*TDG*	*BCCIP*	*CDK7*	*KPNA2*	*TP53BP2*
FANCG	*POLD4*	*UNG*	*BCL2*	*CDK8*	*MAD2L1*	
FEN1	*POLE*	*UNG2*	*BIRC5*	*CDKN1A*	*MAD2L2*	
LIG1	*POLE2*	*WRN*	*CCNB1*	*CDKN1B*	*MCM2*	
LIG3	*POLE3*	*XPA*	*CCNB2*	*CDKN2A*	*MCM3*	
LIG4	*RAD1*	*XPC*	*CCNC*	*CDKN2B*	*MCM4*	
MBD4	*RAD17*	*XRCC1*	*CCND1*	*CDKN3*	*MCM5*	

Genes listed as having association with OVCA in HGMD:

ACACA	*GSTM1*	*NCKIPSD*
ADH1B	*GSTO2*	*NEKB1*
AGER	*GSTP1*	*NOTCH*
AGO2	*GSTT1*	*OGG1*
APOBEC3B	*HFE*	*P14ARF*
ARL11	*HIP1*	*PCM*
ATAD5	*HRAS*	*PGR*
ATF1	*IGF2*	*PPMID*
BPIFC	*IL1A*	*PSMC3IP*
CAV1	*IL23R*	*RB1*
CLTC	*ILI6*	*RNF213*
COL18A1	*ILIR1*	*SLC4A7*
CREBBP	*ITK*	*SLX4*
CYP17A1	*KI*	*SPINK1*
CYP1A1	*KMT5A*	*TERT*
DCC	*KRAS*	*TNFRSF13B*
DCP1B	*LIN28B*	*UGT2A3*
DROSHA	*LPAR6*	*WAS*
E2F2	*LPL*	
EPHX1	*MIR191*	
FMR1	*MIR423*	
FRG1	*MTR*	
FSHR	*MYH9*	

Table 4. Non-Panel genes analyzed for high impact variants Genes were chosen based on their involvement in either DNA repair cell cycle control, or due to being listed as having "disease causing mutations" (DM) associated with OVCA in HGMD.

2.4 In silico SNP assessment

In silico variant assessment of single nucleotide polymorphisms (SNPs) was performed using online available bioinformatics tools, HGMD (Stenson et al., 2003), dbSNP (Sherry et al., 2001), ExAC (*http://exac.broadinstitute.org*), SIFT (Ng, 2003) and PolyPhen (Gnad et al., 2013). Variants were stringently filtered to include only exonic SNPs most likely to have a moderate to high effect impact on protein function (frameshift, nonsense, and missense) while also rare, (<0.02 MAF ExAC; European, non-Finnish) and predicted to be damaging by SIFT and/or PolyPhen, which take into consideration parameters such as amino acid substitution and evolutionary conservation. Variants passing these criteria were confirmed by forward and reverse strand Sanger sequence (GENEWIZ, South Plainfield, NJ) unless otherwise specified. Additional bioinformatic resources were consulted to further describe variants including: dbSNP, UCSC Genome browser, and the Human gene mutation data base (HGMD) which lists variant risk predictions from Mutation predictor risk assessment, Likely hood ratio test for functional prediction (LRT), Mutation taster prediction, and Genomic Evolutionary Rate Profiling (GERP).

2.5 Next generation sequencing validation

Next generation sequencing and other high throughput methodologies are fast and efficient at the cost of accuracy. Therefore, all variants of interest called by WES were validated by Sanger sequencing. Using the primers listed in Table 5, 400-700 bp DNA products, which included the variant of interest were amplified by PCR (Polymerase Chain Reaction) followed by agarose gel electrophoresis to check for amplification and possible contamination. The amplified DNA fragments were then purified by spin column and quantified before being sent out to GENEWIZ for Sanger sequencing.

2.5.1 Primer design

Primers for PCR amplification and targeted sequencing were designed using Primer3Plus *(http://www.bioinformatics.nl/cgi- bin/primer3plus/primer3plus.cgi),* which includes information regarding primer binding specificity, GC percentage, and melting temperature (Table 5).

Gene	dbSNP ID	Variant	Forward Sequencing Primer	Reverse Sequencing Primer
ATM	N/A	c.2503_2507del	AAGAAGAACTTTCATTCTCAGAAGTAG	TTTTTCATCATATAATCCCTATGCTC
ATM	N/A	c.5697_5698insA	TGGTGTACTTGATAGGCATTTGA	TCTGAGCTTTTCCACACTGC
ATM	rs1800054	c.146C>G	GAGCTACAGAACGAAAGGTAGTAAA	TTTCCTCTAATCTGAAGTCTTGTGAA
ATM	rs138327406	c.4388T>G	ATCAGAAAATTCTTCTTGCCATA	CAGGAGGTTGAGGATGCAGT
ATM	rs28904919	c.998C>T	TTTTGTGGGAGCTAGCAGTG	GGTGGCTCATGCCTGTAATC
ATM	rs56009889	c.6919C>T	GTGGGGAGATGTCATGCAG	GGGACACCAATGCCTCTACTT
ATM	rs35203200	c.7618G>A	CCTCAGATAAGAAAAGA	TGCAGTGGGTAGAGCGTG
AXIN1	rs143974067	c.1018G>A	TTCCTGAAGACAAAGCCCAG	GTCACTAACATGCCCTGCTT
BRCA1	rs4986852	c.3119G>A	TCCCATAGGCTGTTCTAAGTTATCTG	CAGAGGCAACGAAACTGGACTCA
BRCA1	rs1800744	c.4535G>T	TTGAGCTATTTTTCTAAAGTGGGCTTA	AGGCAACATGAATCCAGACTTCTAG
BRCA2	rs80358479	c.1889C>T	GCCTCTGAAAGTGGACTGGA	GCAGGCATGACAGAGAATCA
BRCA2	rs28897747	c.8149G>T	TAAAACTAGTAGTGCAGATACCCAAAAAGTG	CAATGACTGATTTTTACCAAGAGTGCAAA
BRCA2	rs11571747	c.8567A>C	TGGAGGAAATGTTGGTTGTGTTGA	CCTTCATGTTCTTCAAATTCCTCCTGA
BRCA2	rs11571833	c.9976A>T	CACCTGTCTCAGCCCAGATGAC	AGTTGTAATTGTGTCCTGCTT
BUB1B	rs28989188	c.1227A>C	TTCCCCACTTTACGCTTTTG	ACCATAGAAGGCAGCAGTGG
CHK1	N/A	c.1564-1565insA	TGAAGTGCCTCTAAAGTTTCCA	TGTTCACACAATGATGAAACCA
CHK2	rs587780185	c.565A>G	ATCACAGTGGCAATGGAACC	CTCCCAAAGTGCTGGGATTA
ERCC6	rs201486862	c.2137A>G	TCGGATCATTCTGTCTGGCT	ATGAGCCTGGCCATCTTTCT
FANCM	rs144567652	c.5713C>T	TCTAGCACTTCAGGGGCATC	TGAAGTGAGCTGTTAGCCATCC
HMMR	rs146791423	c.1054G>T	ACCTCACAATGCCATTCCA	AAGCTGAAAGGCTGGTCAAG
MCM4	N/A	c.1610-1611del	GCGGGACAAGGAAGGATTTT	CATGTTCACGGTGGAGAAGG
MSH6	rs63751005	c.620T>C	TGAACTGGGGCTGGTATTCA	AAGCACACACCATATGCACG
MUTYH	rs34612342	c.494A>G	GTCTCTTTCTGCCTGCCTGT	CTACGTTGCCATCCACCAC
MUTYH	rs36053993	c.1145G>A	AACACTGGACAGTGCCACCT	AAGGGTCAAGGGGTTCAAAT
NBN	rs61754966	c.511A>G	CAACAAAGAAATTTGGGGAAC	GCAGTGACCAAAGACCGACT
PALB2	rs45551636	c.2993G>A	TTTGGCTTAGGGCATTGTTT	GACATGTCTGGCTTCCACCT
PALB2	rs45532440	c.2014G>C	CCTGATGAAGACTTTGGACCTC	TAAGATGGGGAAAGCAGGTG
PALB2	rs200283306	c.3508C>T	TCTGTCTGGACATAAACAAGCAA	ACTCTCAGCGTGGGTGTGAT
PALB2	rs45478192	c.2816T>G	ATCTTTCAGATTCTTTCAAGACTCAAGCC	CTGGATTAAACAAAAATGAAACAACCAAGC
PALB2	rs45494092	c.1010T>C	ATTTCACCAGGGCGACTACA	TTGACTCAAAGGGCTCCACT
PALLD	rs138897963	c.909A>T	ACCTCAGCAGATGTTGTGTC	ATGGGTGCCTAAATGTCGGA
PMS2	rs200513014	c.1004A>G	CAGTGGCTGCTGACTGACAT	GTTGCAGTGAGCTGAGATCG
POLK	N/A	c.1336del	TGAATAGGCTATGGGAGAAAGAA	GGCATTTATTGCAGGGAGTG
POLQ	rs148626322	c.7537C>T	TCCCAAAGAGGGTTACAGGA	AGGCTGAGCGTCAAGCTATC
RAD1	N/A	c.1154del	CGGCCACCTTTAGACTCTTG	TTGGGAGTTCTGAGCAGTGTT
RAD51D	rs587781756	c.511C>T	CCTGCAGCAAAACGTCCTAT	AGTAGGACACCTGCCCACAG
RAD51D	rs387906843	c.616C>T	ACCACTGTGACAACTGACCA	AGTAGGACACCTGCCCACAG
RAD52	rs4987208	c.1245T>G	TAGCAGGAAGCGGAAACG	ACTGCAGTGGGCTCTCAGTC
RAD52	rs4987207	c.806C>A	TCCAGTTCCTCTTTGGTCCT	AGGATCTCCCCTTAATTTTTGTG
REC8	N/A	c.1622G>A	GACCTTCCCCCACTACACAG	TGGGGATGGGAGAAGTAGAA
RECQL	rs150889040	c.962G>A	GAAGCTCTGACCATCCCTGA	CAACAGTTGCCACTACTACCTG
TP53I3	rs145078765	c.755C>G	TCTGAAATCGGGTTCCCTCT	AGGCCTCATAAATGGTGAACTT

Table 5. Primers for the amplification and sequencing of indicated genomic loci

2.5.2 PCR amplification

Polymerase chain reaction (PCR) is a frequently used method for the exponential amplification of short sequences of DNA. PCR requires the use of primers (Table 5) that a thermo-stable DNA polymerase extends to replicate the target (HotStarTaq Plus DNA Polymerase, QIAGEN Fast Cycling PCR KitCat No./ID: 203743). Each round of amplification requires a three-stage temperature fluctuation process to facilitate the reaction (see PCR Protocol below). For ample product yield, the thermocycler (Mastercycler, Eppendorf, Hamburg, Germany) was set for 36 rounds of amplification.

PCR reaction set up	20 µl
Qiagen Master Mix =	10 µl
Primer (forward) =	2 µl
Primer (reverse) =	2 µl
DNA (25ng/µl) =	2 µl
sterile dH^20	4 µl

PCR protocol:

1. Original denaturation: The DNA was heated at 95°C for 5 minutes to denature.

PCR Cycle begins:

2. Denaturation: 96°C for 5 seconds

3. Primer Annealing: 55-62 °C (optimized for specific primer pair efficiency)

4. Elongation: 68 °C for 15 seconds

PCR Cycle ends (new cycle begins)

5. Final elongation: After the last PCR cycle (36 in total), the reaction was incubated at 72 °C for 2 minutes to ensure full extension of any remaining ssDNA.

2.5.3 Agarose gel electrophoresis

Agarose gel electrophoresis is a simple method that separates DNA fragments by molecular weight/size. DNA molecules carry a negative charge due to the phosphate backbone and migrate towards the positively charged end of an electric field. For visualization, ethidium bromide is incorporated into the gel as it intercalates between the bases of dsDNA and fluoresces under UV light. The size of the DNA fragment determines the speed at which it migrates. The larger the fragment, the slower it will migrate through the porous gel. Therefore, DNA product size can be measured by comparison with a standardized ladder. This method is used for the confirmation of PCR product and to identify possible contamination that occurred during the amplification process.

Agarose Gel Electrophoresis Protocol: 1 g of agarose powder (Thermo Scientific, Waltham, MA, Cat# R0491) was dissolved into 100 ml of TAE buffer by boiling. 15µl of ethidium bromide was then mixed into the solution before allowing it to solidify in a gel forming tray. Once solidified, the gel was placed onto an electrophoresis plate, and submerged in TAE buffer. 2µl of loading dye (QIAGEN Fast Cycling PCR KitCat ID: 203743) was added to the PCR product. 2µl of PCR product was loaded in each well. For PCR product size measurement, 2µl of a 100bp DNA ladder was also added. The gel was then given 45 minutes of a 400-mA current at 80 volts, followed by visualization using a UV source (Fisher Biotech, Wembley, Austrailia, Electrophoresis Systems).

2.5.4 Sanger sequencing

Sanger sequencing is a commonly employed method for the accurate sequencing of DNA using chain-terminating (dideoxy) nucleotides incorporated by a DNA polymerase during in-vitro PCR amplification. Each nucleotide version, ddATP, ddTTP, ddCTP, ddGTP, is labeled

with a unique fluorescent dye so that during replication, DNA fragments are randomly terminated with the fluorescent label. The amplified products are then separated by Capillary Array Electrophoresis (CAE), which detects fluorescently labeled fragments and provides an ordered sequence of the fragments based on product length to be computationally assembled. The Sanger method, developed in 1977 by Frederick Sanger, was once the most widely employed sequencing technology. While high throughput sequencing (Next Generation Sequencing) has become the current preferred methodology, Next-Gen results are often validated using Sanger sequencing as it is considered more reliable.

Sanger Sequencing Preparation Protocol: PCR products were purified by spin column (Qiagen QIAquick PCR Purification Kit Cat# 28104) and quantified by spectrophotometer (NanoDrop 2000, Thermo Scientific). PCR reactions were prepared before shipment to GENEWIZ and trace files were visualized using SnapGene Viewer v3.3.3.

Sequence Reaction	15 µl
Primer =	5µl
DNA (25ng/µl) =	2µl
sterile dH^20	8µl

2.6 Cell line assessment and cell culture

HeLa cells were used for all functional experimentation during this project. This cell line was chosen based on various necessary criteria including having at least some wild type (WT) p53 expression (Figure 4A), an intact homologous repair pathway, a doubling capacity of every 24 hours (Figure 4B) and epithelial in origin. While HeLa is a cervical cell line, it is often used in cancer studies for these reasons. Although use of an ovarian cell line is ideal, most immortal ovarian cell lines do not have WT p53 activity (Mullany et al., 2015; Fleury et al., 2015)

meaning processes downstream of p53 such as DNA damage, cannot be assayed accurately. Additionally, the issue of cell line misidentification is common with many "ovarian cell lines" (Korch et al., 2012). All cell lines considered for experimental purposes were sent for authentication (Wayne State University at Applied Genomics Technology Center). The C13 ovarian cell line obtained from an outside lab was confirmed to be ME-180 cervical cells. All other cell lines were authenticated as being correctly labeled.

HeLa cells stably transfected with p.DR-GFP were a kind gift from Dr. Jeffery Parvin of Ohio State University. Cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% Fetal Bovine Serum (HyClone, GE, Little Chalfont, United Kingdom Cat# SH30396.03), NaHCO3 (3.7g/L) 1% penicillin strep (Gibco Life Technologies, Carlsbad, CA Cat# 15140-122) and incubated at 37°C 5% CO_2. For passaging, cells were harvested using trypsin, and split 1:6. Freeze downs of aliquots were maintained in a freezing media of 10% DMSO, 50% DMEM, 40% FBS and kept at -140°C.

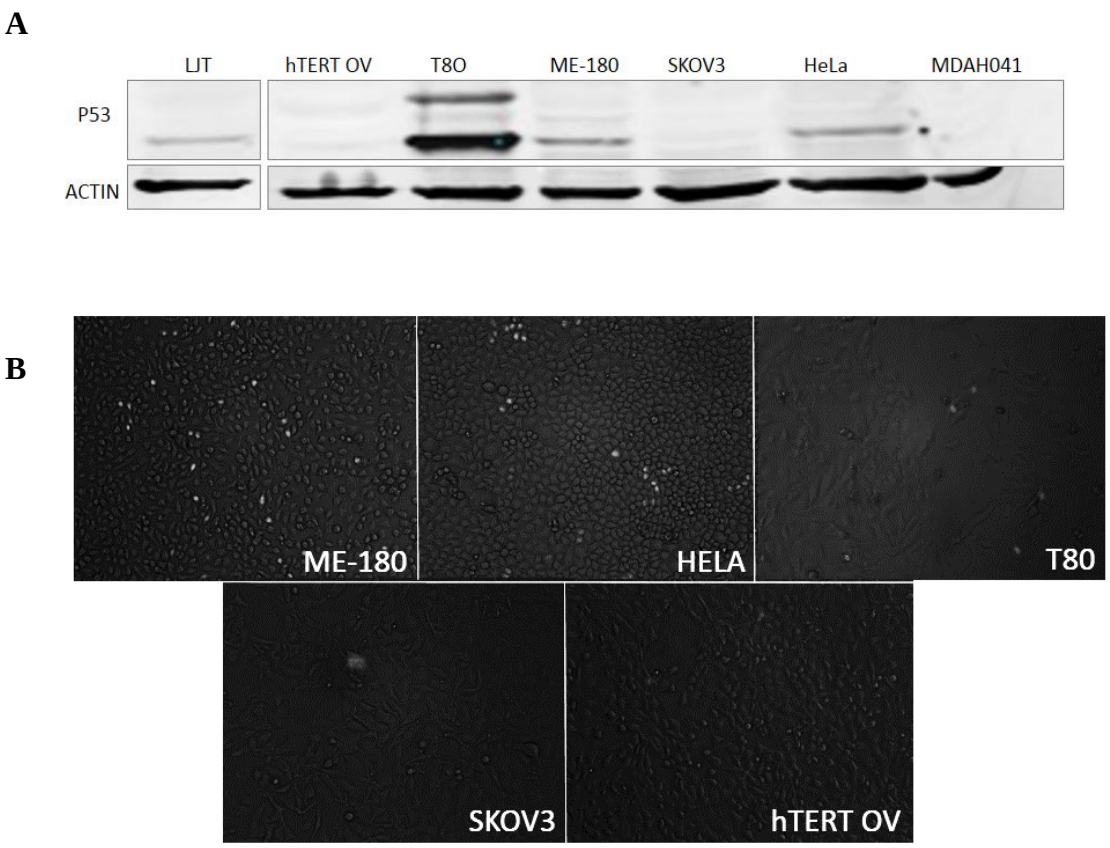


Figure 4. Cell line analysis

A. Western blot for p53 expression in cell lines. p53 expression observed in positive control (LJT), ME-180 (originally thought to be C13 ovarian cell line and shown to be cervical of origin during authentication) and HeLa. p53 expression was not detected in hTERT immortalized ovarian cells (hTERT OV), SKOV3 cells or TP53-negative control MDAH041. Abnormal p53 expression was detected in T80 SV40-Transformed ovarian cells. **B**. Microscopy images of GFP activity as a proxy for intact HRR in ME-180, and HeLa (see section 2.8 Homologous Recombination Repair Assay). Very little GFP activity was observed in T80 cells and none in SKOV3 or hTERT immortalized ovarian cells.

2.7 Protein detection and quantification

Western blot was employed to assay protein expression and quantification. First, cells were lysed, and proteins solubilized using ice cold RIPA (Radio immunoprecipitation assay buffer) supplemented with protease and phosphatase inhibitors. Protein concentration was determined by Bradford assay. Proteins were then reduced and denatured by adding 4x Laemmli buffer plus β-mercaptoethanol to lysates (9:1 ratio respectively), in addition to boiling at 95°C for 5 min using thermocycler. Lysates were stored at -80°C until use for SDS page (sodium dodecyl sulfate polyacrylamide gel electrophoresis).

Next, 50-70μg of whole cell lysate was loaded into each well of a polyacrylamide gel and separated by electrophoresis. Gel percentage varied by size of protein of interest. Stacking gels were prepared at 5%. 7.5% separating gels were used for proteins above 100 kDa and 12% gels for proteins below 100 kDa. Following separation, proteins were transferred from the polyacrylamide gel to a nitrocellulose membrane by electrical current (250 -350 mA). Small molecular weight proteins were transferred for approximately 1 hour at room temperature (RT), large molecular weight proteins 3-4 hours RT or 18 hours overnight at 4°C.

Membranes were then blocked with 5% bovine serum albumin (BSA) in Tris Buffered Saline with Tween 20 (TBST). Next, membranes were incubated with protein specific antibodies at RT for two hours, given three 10 min washings with TBST, incubated with secondary antibody for 1 hr. RT, and followed by three more 10 min TBST washings. Images were obtained by LI-COR Biosciences, Lincoln, NE, Odyssey Blot Imager and protein expression was quantified using ImageJ software and normalized to appropriate loading control (ACTIN for smaller proteins, VINCULIN for large molecular weight proteins).

RIPA lysis buffer	80 ml
1 M Tris-HCL ph. 7.4	4 ml
1.5 M NaCl	8 ml
1% NP40	800 µl
5% Na-deoxycholate	4 ml
.4 M EDTA	200 µl
dH20	63 ml

The day of lysis, 10 µl of protease inhibitor cocktail (PIC), sodium pyrophosphate (NaPP), sodium fluoride (NaF), sodium orthovanadate (NaV), and phenylmethane sulfonyl fluoride (PMSF) was added to every ml of RIPA lysis buffer

SDS page stacking gel	5%
4x stacking buffer	1.25 ml
Acrylamide/Bis 40:32%	.5 ml
dH^20	3.25 ml
10% APS	25µl
Temed	5 µl

SDS page separating gel	7.5%	12%
2x separating buffer	5 ml	5 ml
Acrylamide/Bis 40:32%	2.5 ml	3 ml
dH^20	2.5 ml	2 ml
10% APS	100 µl	100 µl
Temed	25 µl	25 µl

Transfer buffer: LMW <100kDA HMW >100kDA

TRIS	6.06 g	6.06 g
Glycine	28.8 g	28.8 g
Methanol	400 ml	200 ml
dH^20	1600 ml	1800 ml
SDS	-----	.01%

Antibodies for protein detection:

Protein	*Cat#*	*Manufacturer*
ATM	sc-377293	Santa Cruz Biotechnology, Dallas, TX
BRCA1	sc-6954	Santa Cruz Biotechnology, Dallas, TX
BRCA2	sc-28235	Santa Cruz Biotechnology, Dallas, TX
P53	AHO0152	Invitrogen, Carlsbad, CA
CHK1	sc-7898	Santa Cruz Biotechnology, Dallas, TX
CHK2	sc-9064	Santa Cruz Biotechnology, Dallas, TX
TP53I3	sc-16664	Santa Cruz Biotechnology, Dallas, TX
FANCM	AB97905	Abcam, Cambridge, United Kingdom
REC8	sc-15152	Santa Cruz Biotechnology, Dallas, TX
RAD1	sc-166495	Santa Cruz Biotechnology, Dallas, TX
RAD51D	sc-398819	Santa Cruz Biotechnology, Dallas, TX
HMMR	PA5-21105	Invitrogen, Carlsbad, CA
MCM4	sc-28317	Santa Cruz Biotechnology, Dallas, TX
VINCULIN	700062	Invitrogen, Carlsbad, CA
ACTIN	sc-1616	Santa Cruz Biotechnology, Dallas, TX

2.8 Homologous recombination repair assay

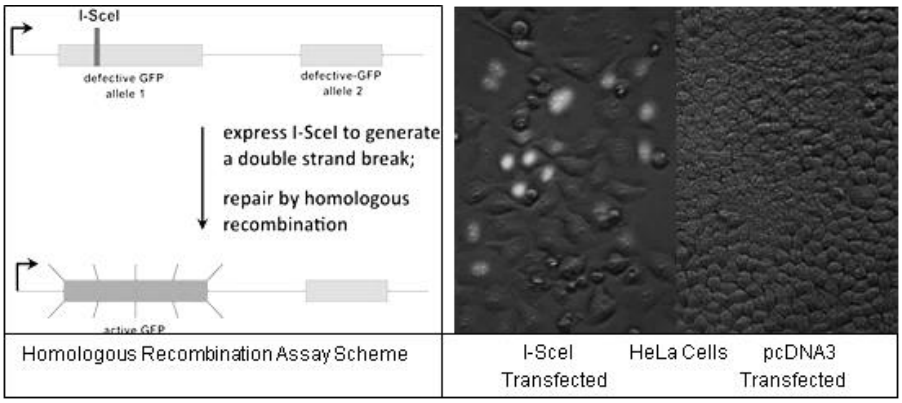
The homologous recombination repair assay is a sensitive method developed to measure Homologous Recombination Repair (HRR) pathway efficiency following double strand breaks (DSB). HeLa cells were stably transfected with pDR-GFP (direct recombinase green fluorescent protein) plasmid and selected with 1.5μg/ml puromycin. The pDR-GFP plasmid contains two inactive GFP alleles; one is inactive due to the presence of an additional sequence that contains the 18bp restriction enzyme recognition site for I-SceI, the other is inactive due to a truncating mutation. When a second plasmid encoding the enzyme I-SceI is transiently transfected into cells containing this plasmid, the I-SceI restriction site is cleaved creating a DSB. The break can only be repaired by the cells' own endogenous HRR pathway using the second inactive GFP allele as a template. In the case of a working HRR pathway, this leads to the restoration and activation of the first GFP allele (Figure 6A). Therefore, the amount of GFP following I-SceI DSB induction proxies for the efficiency of the HRR pathway. This method can accurately assay a gene's involvement in the HRR pathway by measuring the amount of GFP observed after siRNA knock down as compared to a scramble siRNA control.

HRR Assay Protocol: Cells were harvested with trypsin, counted using a hemocytometer and reseeded at 40,000 into each well of a 24 well plate. The next day, media was replaced with 450 μl serum and antibiotic free media and cells were transfected with I-SceI and siRNA specific to the gene of interest using Lipofectamine 3000 (Thermo Fisher, Waltham, MA, Cat# L3000015) according to manufacturer's protocol. As a negative control, cells were transfected with p.cDNA3 empty vector in place of I-SceI to gauge background GFP signal. As a positive control, cells were transfected with scramble siRNA plus I-SceI. Each condition was performed in triplicate. Forty-eight hours post transient transfection of the I-SceI containing plasmid, cells

were visualized for GFP signals using fluorescence microscopy (Olympus 1X71) followed by harvesting and analysis by flow cytometry (BD FACSCanto II and BD FACS Diva Software v8.0.1, BD Biosciences, San Jose, CA). Gating procedure was set to select singlets, live (DAPI) and GFP (FITC) positive cells (Figure 6B). A t-test was employed to compare cells transfected with a siRNA knockdown of a gene of interest to the siRNA scramble positive control and adjusted for background GFP (negative control).

Gene	*Cat#*	*Manufacturer*
Scramble	1027310	Qiagen, Hilden, Germany
ATM	S100299299	Qiagen, Hilden, Germany
BRCA1	SC-29219	Santa Cruz, Dallas, TX
BRCA2	S102653434	Qiagen, Hilden, Germany
CHK1	S100287658	Qiagen, Hilden, Germany
CHK2	S102224264	Qiagen, Hilden, Germany
FANCM	S104158280	Qiagen, Hilden, Germany
HMMR	S102653196	Qiagen, Hilden, Germany
MCM4	S100300818	Qiagen, Hilden, Germany
RAD1	S102653462	Qiagen, Hilden, Germany
RAD51D	S100045094	Qiagen, Hilden, Germany
REC8	AM16708	Ambion (Invitrogen), Carlsbad, CA
TP53I3	S100069636	Qiagen, Hilden, Germany

A.



B.

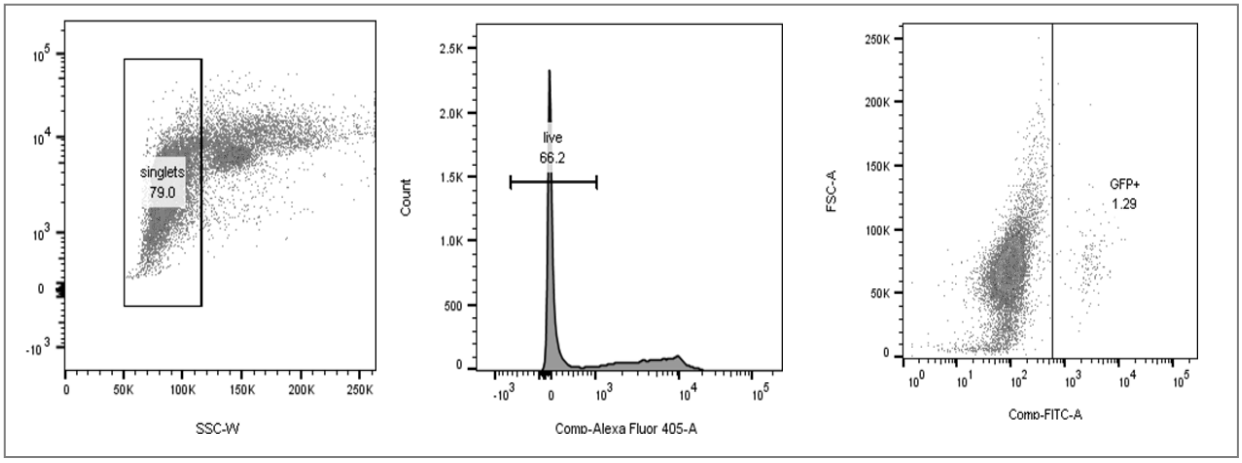


Figure 5. Homologous recombination assay A. Schematic of DSB induction during I-SceI cleavage and GFP induction by HRR. **B.** Flow cytometry gating procedure for selection of GFP positive cells.

2.9 Clonogenic survival assay

The clonogenic survival assay tests the ability of a single cell to form a colony and is a well-established method to determine the importance of a gene to cellular survival as well as to test cell sensitivity to DNA damaging treatments such as ionizing radiation or chemotherapeutic drugs. Certain drugs are of interest due to their ability to create double strand breaks in DNA. Because DSBs are repaired using the HRR pathway, cells efficient in this pathway retain the capacity to repair themselves and produce colonies after exposure, while those deficient in HRR do not**.** Therefore, clonogenic assays involving drug treatments that induce DSBs are an excellent tool for identifying genes important to the HRR process, as well as for indicating which chemotherapeutic drug may be most appropriate for targeted therapy considering the genomic background of certain tumors.

Clonogenic survival assay protocol: Cells were plated (350,000) in 60-mm tissue culture dishes and incubated overnight at 5% CO_2 and 37°C. The next day cells were transfected with siRNA using Oligofectamine (Invitrogen, Carlsbad, CA Cat# 12252-011) according to the manufacturer's protocol. 24 hours after transfection, cells were counted, and 100 to 300 cells were reseeded in triplicate of a 6 well plate and placed back in incubator. The following day (48 hours post siRNA knock down) the media was replaced with serum free media containing a DNA damaging reagent such as Cisplatin, Etoposide, Olaparib, or mock control for the duration and drug concentration optimized for an IC_{50} (50% cell growth inhibition) in HeLa. Cells were rinsed twice before adding back fresh media and incubated at 5% CO_2 and 37°C for 1.5 weeks until colonies had formed (>50 cells per colony). For fixation and staining, medium was removed, and cells washed PBS before adding add 2 ml of acidic acid fixation solution for 5 minutes followed by 2 ml 0.5% crystal violet solution for 2 hours at room temperature. Once the

crystal violet was rinsed off, plates were air-dried at RT for up 24 hours. Colonies were counted both by eye and with a colony counter (GELCOUNT, Oxford Optronix, Abingdon, UK). Plating efficiency (PE) was calculated as the number of colonies formed divided by the number of cells seeded x 100%. Survival after drug treatment was determined by calculating the number of colonies formed divided by the number of cells seeded x 100%, adjusted for PE.

2.10 Statistical analysis of functional data

Statistical analyses of homologous recombination repair and clonogenic assays were carried out using R statistical software (version 3.4.1). Prior to the analysis, we checked for and did not find outliers (i.e., data points greater than 3 standard deviations from the mean for each variable). We next investigated the distribution of replicates for each tested condition (reduction in HRR efficiency, plating efficiency, and adjusted survival rate after drug treatment following siRNA knockdown). First, we calculated Shapiro-Wilk's tests, which tests the null hypothesis that a sample distribution was drawn from a normally distributed population. Next, we assessed skewness and kurtosis for each gene per condition. For small samples ($n < 50$), z-scores less than 1.96 for either skewness or kurtosis suggests a normal distribution[14]. We performed approximately six replicates for each condition to reduce inflation of Type II error. The overall pattern of results generated from the HRR and clonogenic assays indicated normal distributions (detailed results available upon request). Paired sample t-tests were then conducted to identify mean differences in the survival rates for each siRNA knock down condition and its respective scramble siRNA control.

CHAPTER 3: RESULTS PART I

3.1 Clinically actionable variants

We performed WES on blood DNA from 48 women with a personal history of OVCA and determined to be at high risk for inheritance of a germline predisposition mutation, but with no known deleterious mutations in *BRCA1/BRCA2*. In total, five clearly pathogenic loss of function variants were identified (Table 6). Four were in genes currently featured on newer comprehensive HBOC panels; two novel frameshift variants in *ATM* (c.2503_2507del and c.5697_5698insA) and two truncating variants in *RAD51D (*rs587781756 p.Q171* and rs387906843 p.R206*, as well as a pathogenic variant in a non-panel gene, *FANCM* (rs144567652 p.R1931*) previously found to be strongly associated with hereditary risk of breast cancer (Peterlongo et al., 2015). Pathogenic variants in genes with an associated cancer risk are considered clinically actionable, meaning a medical intervention, or risk reduction measures are available.

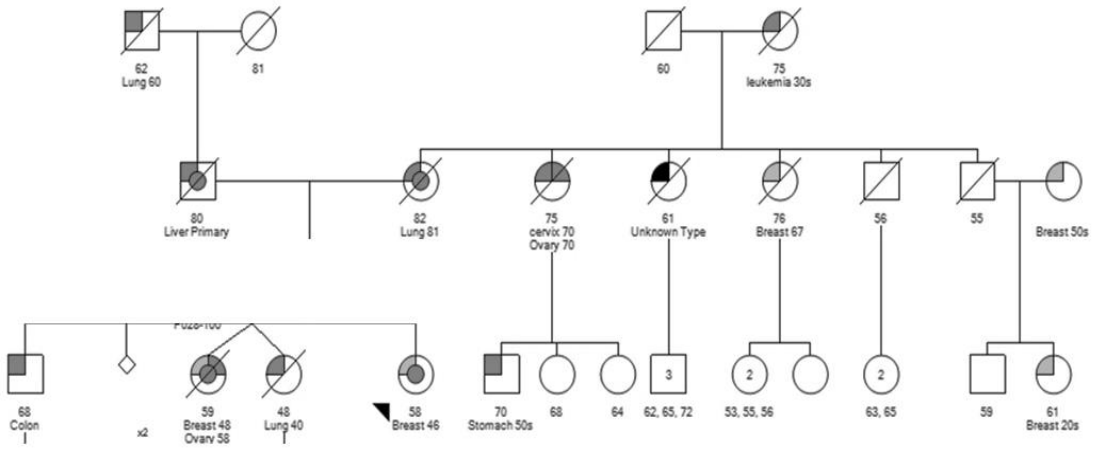
Table 6. Clinically Actionable Variants

ID	Gene	Consequence	AA	dbSNP ID	Variant	MAF	OBS
OCF28-1	ATM	FRAMESHIFT	CATCTG>C	N/A	c.2503_2507del	N/A	1
OCL56	ATM	FRAMESHIFT	G>GA	N/A	c.5697_5698insA	N/A	1
OCJ19	FANCM	STOP	R1931*	rs144567652	c.5713C>T	0.0009	1
OCH26	RAD51D	STOP	Q171*	rs587781756	c.511C>T	N/A	1
OCK1	RAD51D	STOP	R206*	rs387906843	c.616C>T	0.0001	1

Table 6. Clinically Actionable Variants are those of high impact (frameshift or stop gain) in genes already associated with either breast or ovarian cancer. AA= Amino acid change, MAF= Minor allele frequency (ExAC, European non-Finnish) OBS= Number of times variant was observed in sample. N/A= Not Available. All variants listed were confirmed by Sanger DNA Sequencing.

ATM (Ataxia Telangiectasia Mutated) codes for a protein kinase important for DNA damage recognition and activation of substrates including p53, BRCA1, and other homologous recombination repair factors. Homozygous mutations in *ATM* cause ataxia-telangiectasia, a rare inherited autosomal recessive disorder which affects the immune and nervous systems and leads to increased sensitivity to radiation and cancer susceptibility. Although heterozygous *ATM* mutation carriers do not have ataxia-telangiectasia, they have a 17-52% lifetime risk of developing breast cancer (Broeks et al., 2000). However, despite association of *ATM* with ovarian cancer in recent literature, carriers are not routinely counseled with this information as exact risks are unknown. One patient with an *ATM* pathogenic variant in our sample (OCF28-1) had a family history of liver, lung (n=2) and breast cancer, on the same parental side of the family. The proband herself was first diagnosed with breast cancer at the age of 48 before a secondary diagnosis of OVCA at 57 (Figure 6A). The second carrier of an *ATM* frameshift mutation (OCL56) was diagnosed at 73 and had a family history of OVCA (two additional cases besides herself) as well as two cases of breast cancer, all on the maternal side (Figure 6B).

A. Kindred OCF28-1



B. Kindred OCL56

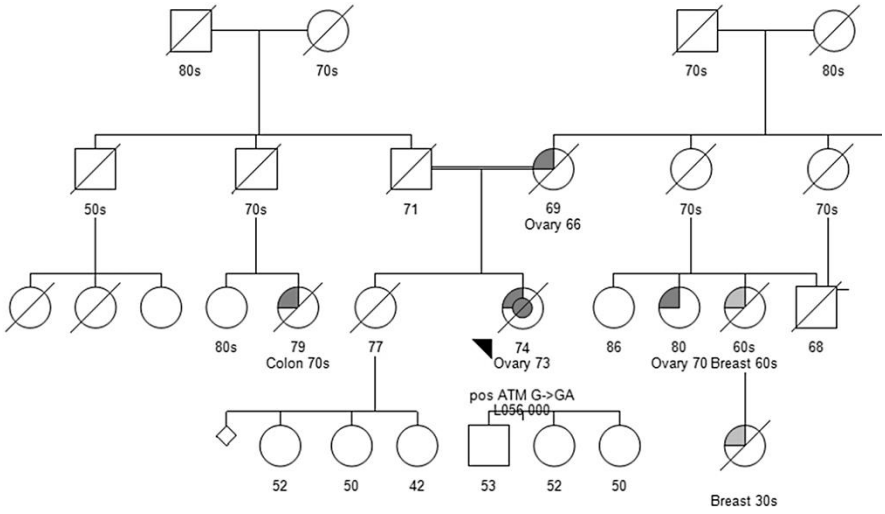


Figure 6. ATM LOF variants may lead to increased OVCA risk A. Kindred OCF28-1. Proband (indicated with arrow) is positive for *ATM* c.2503_2507del pathogenic variant. **B.** Kindred OCL56. Proband (indicated with arrow) is positive for *ATM* c.5697_5698insA pathogenic variant.

The second gene featuring pathogenic variants in our sample, *RAD51D* (paralog of *RAD51*), has recently been identified as a moderately penetrant gene in hereditary ovarian cancer (Janatova et al., 2015; Thompson et al., 2013; Meindl et al., 2010). RAD51D forms a complex with RAD51B, RAD51C and XRCC2 in order to bind single stranded DNA, a necessary process for DNA repair by homologous recombination and is required for RAD51 foci formation upon DNA damage induction (Tarsounas, Davies, and West, 2004). Although rare among familial breast cancer patients (Thompson et al., 2013), loss of function variants in *RAD51D* have been associated with a relative risk for OVCA of 6.30. (95% CI 2.86-13.85) (Loveday et al., 2016). Two pathogenic nonsense SNPs in *RAD51D* were discovered in our sample. One carrier (OCH26) was diagnosed at the age of 61 and had a family history of prostate (n=2), breast (n=2) and ovarian cancer on her paternal side, while the second carrier (OCK1), diagnosed at 67, had a comparatively weak family history with a single diagnosis of colon cancer on her paternal side and lung cancer in a maternal aunt.

In addition, a pathogenic nonsense mutation in a non-panel gene, *FANCM* (rs144567652, p.R1931*) was identified. This variant has been recently associated with increased risk of breast cancer (OR of 3.93) (Peterlongo et al., 2015), warranting contact for further counseling. *FANCM* is the most highly conserved member of the Fanconi Anemia Complementation Group (Schwab et al., 2015). This group is associated with the autosomal recessive genetic disorder, Fanconi Anemia, which is characterized by genomic instability, hypersensitivity to DNA damage induced by crosslinking agents and substantial increased risk of leukemia and other cancers (Bogliolo and Surrallés, 2015). Other members of the Fanconi Anemia complementation group include breast and ovarian cancer-associated genes; *RAD51C (FANCO), BRCA1 (FANCS), BRCA2 (FANCD1), BRIP1 (FANCJ) and PALB2 (FANCN).* *FANCM* encodes for an ATP-

dependent helicase important for the resolution of DNA: RNA hybrids, thus ensuring stability with genome duplication (Schwab et al., 2015). The nonsense variant identified here has been shown to affect protein function by also inducing exon skipping (Peterlongo et al., 2015). The carrier (OCJ19) of *FANCM* rs144567652 was diagnosed with OVCA at 49 years of age and had a family history of breast (n=2), multiple myeloma, leukemia, and ovarian, all on the maternal side of her family (Figure 7).

Kindred OCJ19

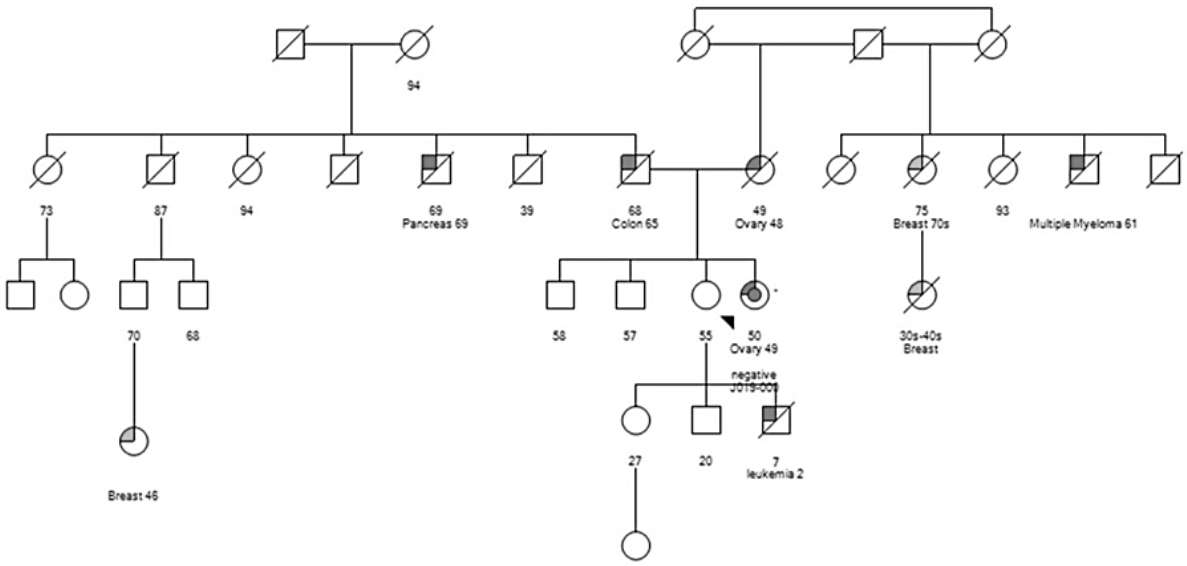


Figure 7. Kindred OCJ19. Proband is carrier of *FANCM* pathogenic variant

3.2 Variants of unknown clinical significance detected in HBOC panel genes

As most women in our sample were not found to be carriers of a clearly pathogenic mutation upon WES, I next sought to identify potentially deleterious variants in HBOC panel genes and found that 23 women in the sample (37%) harbored one or more rare and predicted to be damaging variants of unknown significance (VUS), in panel genes; *ATM*, *BRCA1, BRCA2, CHK2, MHS6, MUTHY, NBN, PALB2,* and *PMS2* (Table 7). Analysis of the Human Gene Mutation Database (HGMD) (Stenson et al., 2003) revealed that six of these variants are annotated as "disease causing" (DM), ten as "possibly disease causing" (DM?) and three as "disease associated polymorphism" (DP) in the Human Gene Mutation Database (HGMD) (Stenson et al., 2003).

Carriers of deleterious variants in either *ATM, CHK2*, *PALB2* or *NBN* are typically counseled for their risk of breast cancer, but not ovarian cancer despite associations in current literature (Walsh et al., 2011; Thorstenson et al., 2003; Norquist et al., 2015; Lawrenson et al., 2015). In addition to the pathogenic *ATM* frameshift mutations previously discussed, we detected an additional six rare and predicted to be damaging missense VUSs in *ATM;* rs1800054, rs138327406, rs28904919, rs1801673, rs56009889, rs35203200. The *ATM* variant rs1800054 (p.S49C) has recently been implicated as associated with a slightly increased risk for breast cancer (OR 1.08 (C.I .95-1.22) for heterozygotes, 1.44 (.39-5.32) for homozygotes (Fletcher et al., 2010). *ATM* variant rs138327406 (p.F1463C MAF= 0.002) is listed as a disease-causing mutation in HGMD and was found in three of six women of Ashkenazi Jewish (AJ) descent in our sample, always in combination with a second rare polymorphism 266 amino acids apart (rs2227922, p.P604S, MAF=0.003) which was predicted to be benign. These variants were not seen in any other women in our sample and linkage data suggests they are not in disequilibrium

(r^2 =0.5, Haploreg v4, CEU). Therefore, it is suspected that there may be a founder effect resulting in the coupled segregation on a single haploblock in the AJ population. One participant (OCG29) was found to have inherited both variants on the same parental allele. We were not able to confirm co-segregation in the other two participants as fresh peripheral blood samples were not available to prepare RNA for this analysis. However, we did find that the unaffected daughter of OCD16 was wild type for both variants, suggesting likely co-segregation.

Similarly, with *PALB2*, a detected a pair of rare SNPs inherited together, rs45532440 (p.E672Q MAF= 0.02) and rs45551636 (p.G998E MAF= 0.02) $r^{2=}$0.69, in two unrelated individuals (OCH26 and OCE17-2). PALB2 (partner and localizer of BRCA2), physically interacts with BRCA2, and is critical for the localization and stability of BRCA2 in the nucleus. Females with monoallelic germline loss of *PALB2* have a 2-4 fold increased breast cancer risk (Rahman et al., 2010; Erkko et al., 2007). *CHK2* and *NBN* are also known breast cancer-associated genes in which we found an interesting VUS in our sample. Female *CHK2* and *NBN* pathogenic mutation carriers are at an increased lifetime risk of developing breast cancer with a 2-fold for CHK2 and 3-fold for NBN carriers, (The CHEK2 Breast Cancer Case-Control Consortium 2004). Both patients had a family history of breast cancer and the carrier of *CHK2* had a secondary diagnosis of breast cancer. The p.I232V (rs587780185) variant in *CHK2* is extremely rare (MAF= 00001). SIFT and PolyPhen predict this alteration as deleterious and probably damaging. *NBN* p.I171V (rs61754966) has contradictory annotations among various bioinformatics assessment tools, but is annotated as a disease-causing mutation in HGMD.

Numerous potentially deleterious VUSs in Lynch syndrome and familial adenomatous polyposis associated genes were detected in the study sample. Lynch syndrome (hereditary nonpolyposis colorectal cancer), is an autosomal dominant inherited disorder caused by

mutations in mismatch repair genes; *MLH1*, *MSH2*, *MSH6*, *PMS2*, or *EPCAM* which lead to high risk of colorectal cancer (80% lifetime risk) among others, including cancer of the ovaries (10-15% lifetime risk) and endometrium (71% lifetime risk) (Strafford, 2012). Skin cancer, in the form of Muir-Torre syndrome (a variant of Lynch) is another non-colonic phenotype observed in some Lynch families (Bansidhar, 2012; South et al., 2008). A rare (MAF=0.007), highly conserved (GERP=5.35) and predicted as deleterious VUS was found in the Lynch associated gene, *MHS6* (p.V509A rs63751005). The carrier of this SNP (OCD16) was diagnosed with OVCA at the age of 25, followed by a secondary diagnosis of colon cancer at the age of 65 and had a family history of colon and skin cancer as well. Two patients in the sample were heterozygous for very rare missense *MUTHY* mutations considered to be pathogenic and the cause of MYH-associated polyposis (MAP) in homozygote carriers (rs34612342 p.Y179C MAF= 0.002 and rs36053993 p.G396D MAF= 0.003). Although it is possible that a second pathogenic *MUTHY* variant occurred sporadically in the other parental allele, tissue was unavailable to detect this change in these patients. Biallelic mutations in *MUTYH* have been shown to mimic Lynch syndrome by disrupting base excision repair and resulting in a somatic loss of function of mismatch repair (Morak et al., 2014). The carrier of the *MUTYH* variant, rs34612342, (OCE17-2) had a family history of skin and breast cancer and was a carrier of an additional VUS in the Lynch gene *PMS2*. The carrier of *MUTHY* rs36053993 (OCQ15) was also diagnosed with melanoma and had a family history of colon (n=2) skin (n=2) and ovarian cancer.

Another conspicuous finding in our sample was the occurrence of a specific *BRCA2* truncating mutation in four unrelated individuals. The *BRCA2* variant p.K3326* (rs11571833) results in a 93-amino acid truncation and has a minor allele frequency (MAF) of 0.009 (EXAC non-Finnish). The odds ratio of observing this mutation in our sample relative to its MAF in the

ExAC cohort is 4.95 (Fisher's Exact test p-value = 0.01). It is worth noting that this allele is much more frequent in the Finnish population (MAF=.01). However, even using this more frequent MAF as a reference, our test indicates that the allele is still significantly overrepresented (p = 0.03, OR = 3.71). Although the role of *BRCA2* has been established in breast and ovarian cancer, the K3326* variant is considered to be benign by commercial testing and therefore was not identified in the initial *BRCA1/BRCA2* screening. However, recent literature is in disagreement with this classification and established that this SNP is a risk factor for lung, oral and pancreatic cancers (Akbari et al., 2008; Martin et al., 2005; Rudd et al., 2006) all of which were observed in the family histories of the four K3326* carriers; throat (OCP36), lung (OCK1 and OCF28-1) pancreatic (OCN22), and esophageal cancer (OCN22). The accepted risk for breast cancer in carriers of this SNP is low but significant (p = 0.047, OR 1.53, 95% CI 1.00-2.34) (Thompson et al. 2015). Two of the four carriers had a family history of breast cancer, one of which had a primary diagnosis of breast cancer prior to ovarian cancer. Furthermore, analysis of the GAME-ON database (>15000 OVCA cases and >30,000 controls) indicates that this SNP is also associated with OVCA with a p-value of $2.7x10^{-4}$ and OR (95% CI) = 1.31 (1.22-9.32) for all histologies, and for 8,864 invasive serous OVCA cases versus controls, the p-value was $7.11x10^{-8}$ and OR (95% CI) = 1.57 (1.44-1.70). This data was provided by the Ovarian Cancer Association Consortium (OCAC) (*http://apps.ccge.medschl.cam.ac.uk/consortia/ocac/*). These findings indicate that *BRCA2* K3326* is likely a low risk allele in ovarian cancer.

ID	Gene	Consequence	AA	dbSNP ID	Variant	MAF	OBS	SIFT	PolyPhen	M.PRED	LRT	M.T	GERP	HGMD
OCD13	ATM	MISSENSE	S49C	rs1800054	c.146C>G	0.011	1	DEL	Possibly Damaging	N/A	NEUT	DC	4.22	DM?
OCD16	ATM	MISSENSE	F1463C	rs138327406	c.4388T>G	0.002	3	TOL	Probably Damaging	MEDIUM	DEL	DC	5.56	DM
OCG29	ATM	MISSENSE	F1463C	rs138327406	c.4388T>G	0.002	3	DEL	Probably Damaging	MEDIUM	DEL	DC	5.56	DM
OCK1	ATM	MISSENSE	F1463C	rs138327406	c.4388T>G	0.002	3	DEL	Probably Damaging	MEDIUM	DEL	DC	5.56	DM
OCM13	ATM	MISSENSE	S333F	rs28904919	c.998C>T	0.004	1	DEL	Benign	N/A	NEUT	PM	3.85	DM?
OCJ10	ATM	MISSENSE	N1853V	rs1801673	c.5558A>T	0.006	1	N/A	Benign	MEDIUM	DEL	DC	5.53	DP
OCP43	ATM	MISSENSE	L2307F	rs56009889	c.6919C>T	0.0019	1	DEL	Probably Damaging	N/A	DEL	PM	5.58	DM?
OCP43	ATM	MISSENSE	V2540I	rs35203200	c.7618G>A	0.00005	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A
OCL56	BRCA1	MISSENSE	S1040N	rs4986852	c.3119G>A	0.019	2	TOL	Possibly Damaging	N/A	NEUT	PM	2.01	DM?
OCN45	BRCA1	MISSENSE	S1040N	rs4986852	c.3119G>A	0.019	2	N/A	Possibly Damaging	N/A	NEUT	PM	2.01	DM?
OCD16	BRCA1	MISSENSE	S1533I	rs1800744	c.4535G>T	0.003	1	N/A	Benign	N/A	NEUT	PM	3.98	DM?
OCD16	BRCA2	MISSENSE	T630I	rs80358479	c.1889C>T	0.001	1	DEL	Benign	N/A	N/A	N/A	N/A	N/A
OCN45	BRCA2	MISSENSE	A2717S	rs28897747	c.8149G>T	0.001	1	TOL	Probably Damaging	MEDIUM	DEL	PM	4.5	DM?
OCP9	BRCA2	MISSENSE	E2856A	rs11571747	c.8567A>C	0.001	1	TOL	Probably Damaging	N/A	NEUT	PM	5.28	DM?
OCF28-1	BRCA2	STOP	K3326*	rs11571833	c.9976A>T	0.009	4	N/A	N/A	N/A	N/A	N/A	N/A	DP
OCK1	BRCA2	STOP	K3326*	rs11571833	c.9976A>T	0.009	4	N/A	N/A	N/A	N/A	N/A	N/A	DP
OCN22	BRCA2	STOP	K3326*	rs11571833	c.9976A>T	0.009	4	N/A	N/A	N/A	N/A	N/A	N/A	DP
OCP36	BRCA2	STOP	K3326*	rs11571833	c.9976A>T	0.009	4	N/A	N/A	N/A	N/A	N/A	N/A	DP
OCE27	CHK2	MISSENSE	I232V	rs587780185	c.565A>G	0.00001	1	DEL	Probably Damaging	N/A	N/A	N/A	N/A	N/A
OCD16	MSH6	MISSENSE	V509A	rs63751005	c.620T>C	0.001	1	DEL	Probably Damaging	HIGH	DEL	DC	5.35	DM?
OCE17-2	MUTYH	MISSENSE	Y179C	rs34612342	c.494A>G	0.002	1	DEL	Probably Damaging	VERY HIGH	DEL	DC	5.01	DM
OCQ15	MUTYH	MISSENSE	G396D	rs36053993	c.1145G>A	0.003	1	DEL	Probably Damaging	VERY HIGH	DEL	DC	5.4	DM
OCN71	NBN	MISSENSE	I171V	rs61754966	c.511A>G	.001	1	TOL	Probably Damaging	LOW	NEUT	DC	4.81	DM
OCH30	PALB2	MISSENSE	G998E	rs45551636	c.2993G>A	0.021	2	DEL	Probably Damaging	LOW	DEL	DC	5.84	DP
OCE17-2	PALB2	MISSENSE	G998E	rs45551636	c.2993G>A	0.021	2	DEL	Probably Damaging	LOW	DEL	DC	5.84	DP
OCE17-2	PALB2	MISSENSE	E672Q	rs45532440	c.2014G>C	0.029	3	TOL	Benign	N/A	NEUT	PM	1.83	DM?
OCH26	PALB2	MISSENSE	E672Q	rs45532440	c.2014G>C	0.029	3	TOL	Benign	N/A	NEUT	PM	1.83	DM?
OCH30	PALB2	MISSENSE	E672Q	rs45532440	c.2014G>C	0.029	3	TOL	Benign	N/A	NEUT	PM	1.83	DM?
OCJ3	PALB2	MISSENSE	H1170Y	rs200283306	c.3508C>T	0.0001	1	TOL	Possibly Damaging	N/A	N/A	N/A	N/A	N/A
OCN2	PALB2	MISSENSE	L939W	rs45478192	c.2816T>G	0.0015	1	DEL	Probably Damaging	N/A	DEL	DC	5.81	DM
OCJ19	PALB2	MISSENSE	L337S	rs45494092	c.1010T>C	0.019	1	DEL	Benign	N/A	NEUT	PM	2.61	DM?
OCE17-2	PMS2	MISSENSE	N335S	rs200513014	c.1004A>G	0.0004	1	DEL	Probably Damaging	N/A	DEL	DC	5.73	DM

Table 7. Rare and predicted to be deleterious/damaging variants of unknown clinical significance in sample. AA= Amino acid change, MAF= Minor allele frequency (ExAC, European non-Finnish) OBS= Number of times variant was observed in sample. MUT PRED= Mutation predictor risk assessment, LRT= Likely hood Ratio Test for functional predicting of mutation, DEL= deleterious, TOL= Tolerated NEUT= Neutral, MUT TAST= Mutation Taster

prediction, DC= Probably Disease Causing, PM= Probably Polymorphism, GERP= Genomic Evolutionary Rate Profiling, a score above 2 indicates a highly constrained sequence, HGMD Variant class; DM= Disease causing mutation, DM?= Possible disease causing mutation, DP= Disease associated mutation, N/A= Not Available. All variants listed were confirmed by Sanger DNA Sequencing.

3.3 High impact mutations in non-panel candidate genes

A portion of the missing heritability in OVCA is likely due to risk factors in genes not currently featured on testing panels. The implication of even a highly penetrant mutation would be difficult to interpret if rare, even in a mechanistically relevant gene not previously associated with the disease. Despite selecting for patients with high risk of genetic inheritance, half of the subjects in our sample were not found to harbor a pathogenic variant, nor a variant of unknown significance in any of the 24 panel genes currently tested in HBOC syndromes (Figure 8). I therefore sought to identify rare (MAF ≤ 0.02) mutations in our sample of high functional impact (frameshift or stop gain) in candidate genes. Using DAVID (*https://david.ncifcrf.gov/*) functional annotation and literary searches I compiled a candidate gene list including 115 genes involved in DNA repair and/or cell cycle control, the two pathways most commonly associated with HBOC, in addition to 64 genes having a disease-causing variant (DM) in HGMD for ovarian cancer. A full list of non- panel candidate genes analyzed is provided in methodology section (Table 4).

This analysis uncovered 11 high impact mutations in four cell cycle control genes, *CHK1, RAD1, TP53I3* (n=2), *MCM4,* and six DNA repair genes, *FANCM*, *HMMR, POLK, POLQ, RAD52* (n=2), and *REC8* (Table 8). Importantly, this analysis resulted in the discovery of a clinically actionable pathogenic nonsense variant in *FANCM* (rs144567652) previously discussed. Most of these non-panel genes are not featured in HGMD, and are they are not analyzed during clinical testing. Therefore, I have provided the mouse phenotype seen in knock-out studies where possible. A common phenotypic presentation of many known cancer predisposition genes, such as *BRCA1/2*, includes embryonic lethality in homozygote knockouts

and increased cancer incidence in heterozygotes, which are reported in mouse model studies of some of these genes (Table 8).

The frameshift mutation in *CHK1* (Checkpoint Kinase 1) is notable because much like panel gene *CHK2*, it encodes for a serine/threonine protein kinase required for checkpoint-mediated cell cycle arrest and activation of DNA repair HRR. I also discovered a frameshift variant in *RAD1,* a gene whose protein product functions as part of the 9-1-1 cell cycle checkpoint complex to arrest cellular proliferation in the presence of incomplete DNA replication or damaged DNA, as well as in *MCM4* (Mini-chromosome maintenance complex component 4), a highly conserved helicase protein required for genome replication by initiation of replication fork formation (Sheu et al., 2014). The *TP53I3* (TP53 inducible protein 3) nonsense SNP (rs145078765 p. S252* MAF= 0.0009) is also of great interest as it was observed in two unrelated individuals in our sample. TP53I3 is an oxidoreductase-like protein and an inducer of reactive oxygen species (ROS), that is transcriptionally activated by the tumor suppressor P53 and likely to be involved in P53-mediated apoptosis (Zhang et al., 2015).

Among DNA repair genes, I observed high impact mutations in those encoding DNA polymerases, *POLK* (c.1336del), a translesion polymerase that initiates the continuation of replication through DNA lesions in damaged DNA, and *POLQ* (p.Q2513* rs148626322), a gene associated with micro homology-mediated end-joining pathway (MMEJ), both in the same patient. I also identified truncating variants in chromatid cohesion *REC8*, whose protein product binds sister chromatids during meiosis, and *HMMR* (hyaluronan mediated mobility receptor), which encodes for a cell motility protein that forms a complex with tumor suppressors *BRCA1* and *BRCA2*. Common missense variations in *HMMR* have been shown to modify the penetrance of breast cancer risk in *BRCA1* pathogenic mutation carriers (Maxwell et al., 2011).

Furthermore, two *RAD52* truncating SNPs; rs4987207 p.S346* and rs4987208 p.Y415* were discovered. *RAD52* mediates complementary ssDNA annealing and recruits RAD51 recombinase to promote recombination and HRR. However, the *RAD*52 truncating variants observed in our sample had previously been found to lack an association with OVCA or breast cancer (J. Han et al., 2002).

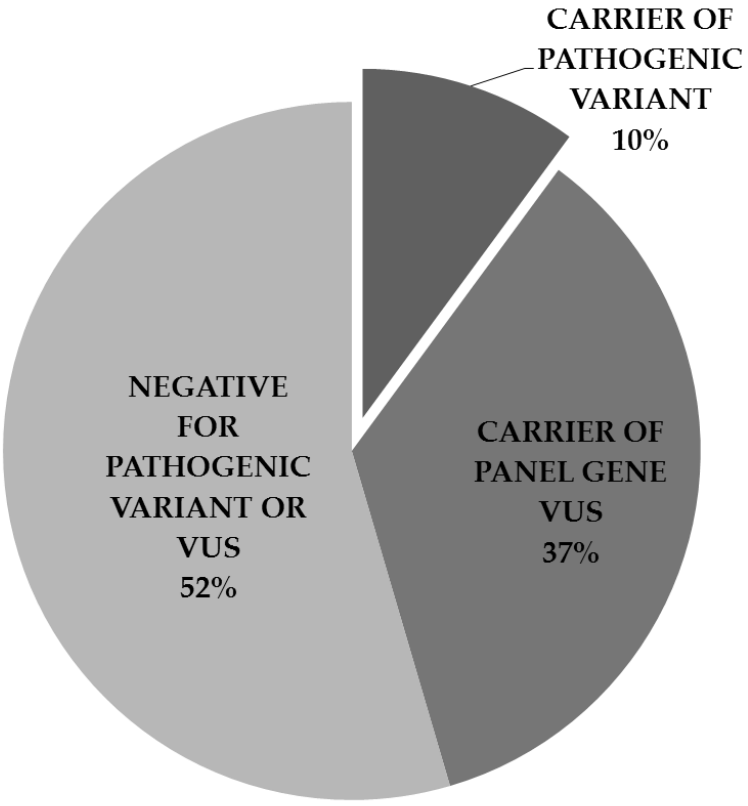


Figure 8. Summary of variant findings amongst our 48 subjects of high risk for genetic inheritance of OVCA.

ID	Gene	Consequence	dbSNP ID	Variant	MAF	Mouse Phenotype +/-	Mouse Phenotype -/-	OVARY RPKM*
OCL60	CHK1	FRAMESHIFT	N/A	c.1564-1565insA	N/A	enhanced tumorigenesis of WNT-1 transgenic mice (Q. Liu et al., 2000)	embryonic lethal (Q. Liu et al., 2000)	1
OCJ19	FANCM	STOP	rs144567652	c.5713C>T	0.0009	None (Bakker et al., 2009)	reduced life span increased cancer incidence (Bakker et al., 2009)	1.1
OCK1	HMMR	STOP	rs146791423	c.1054G>T	0.0035	None (Pujana et al. 2007)	impaired ovarian folliculogenesis (Pujana et al., 2007)	.1
OCN37	MCM4	FRAMESHIFT	N/A	c.1610-1611del	N/A	mammary adenocarcinomas in 80% of females (Shima et al. 2007)	preimplantation and embryonic lethal (Shima et al., 2007)	3.8
OCG24	POLK	FRAMESHIFT	N/A	c.1336del	0.0006	none (Stancel et al. 2009)	spontaneous mutator (Stancel et al., 2009)	6.2
OCG24	POLQ	STOP	rs148626322	c.7537C>T	0.0002	none (Yousefzadeh and Wood 2013)	increased chromosome breaks in peripheral erythrocytes (Yousefzadeh and Wood, 2013)	8.2
OCG23	RAD1	FRAMESHIFT	N/A	c.1154del	N/A	larger, more numerous, earlier onset skin tumors with DMBA-TPA treatment (L. Han et al., 2010)	embryonic lethal (L., Han et al. 2010)	3.4
OCL11	RAD52	STOP	rs4987208	c.1245T>G	0.019	None (Rijkers et al., 1998)	None (Rijkers et al., 1998)	7.1
OCL60	RAD52	STOP	rs4987207	c.806C>A	0.012			
OCL56	REC8	STOP	N/A	c.1622G>A	N/A	None (Xu et al., 2005)	sub-Mendelian frequencies and failure to thrive (Xu et al., 2005)	3.3
OCG14	TP53I3	STOP	rs145078765	c.755C>G	0.001	No knockout mouse found in the literature for this gene.	No knockout mouse found in the literature for this gene.	3.2
OCJ19	TP53I3	STOP	rs145078765	c.755C>G	0.001			

Table 8. High Impact Mutations in DNA repair and Cell Cycle Control Genes, not Featured on HBOC Testing Panels. Rare and high impact variants (frameshift or stop gain) in sample found in DNA repair or cell cycle control genes not currently known to associate with breast or ovarian cancer. MAF=Minor Allele Frequency in Non-Finish Europeans (ExAC). Mouse Phenotype= Available phenotypic information on homozygote (-/-) or heterozygote (+/-) mouse knock out models. Ovary expression data RPKM (reads per kilobase per million) obtained by https://gtexportal.org. *For reference, OVCA genes *BRCA1*=.6, *BRCA2*= .095, *RAD51D*= 4. All variants listed were confirmed by Sanger DNA Sequencing.

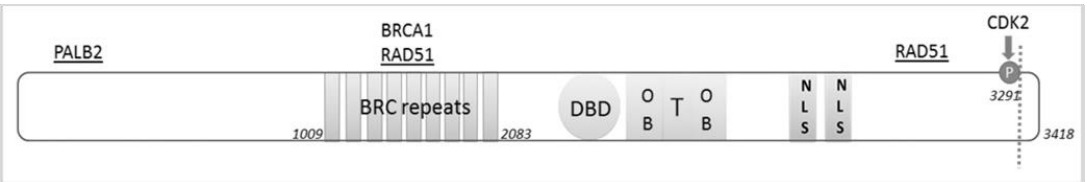
3.4 *BRCA2* p.K3326* truncation as a possible modifier of penetrance

The prevalence of the *BRCA2* K3326* variant (Figure 9A) in our sample, along with the evidence of an association with lung, aero digestive, and pancreatic cancer (Akbari et al. 2008; Martin et al. 2005; Rudd et al. 2006) indicate that this variant may be of minimal risk when inherited alone but could act as a modifier of penetrance to a secondary more deleterious mutation. A portion of the missing heritability in OVCA is likely due to this type of polygenic inheritance. This possibility led to the investigation of other putative pathogenic variants that each of the four carriers had inherited in addition to *BRCA2* p.K3326* (Table 4). I therefore looked for additional rare, and moderate or high impact variants in either HBOC panel genes or candidate genes (involved in DNA repair/cell cycle and with cancer associations in HGMD) amongst the four *BRCA2* K3326* carriers. A complete list of rare and predicted to be damaging variants of moderate impact in cell cycle and DNA repair genes is available in the appendix (Table 12).

In analyzing relevant candidate genes, I identified two patients who along with K3326* were carriers of an additional, clearly pathogenic variant; a *RAD51D* nonsense mutation (OCK1) and *ATM* frameshift mutation (OCF28-1). This observation is interesting because BRCA2 interacts with the RAD51 paralogs and a *BRCA2/RAD51D* double knockdown leads to a greater loss of cellular viability (Jensen et al., 2013). The carrier of both the *ATM* frameshift and *BRCA2* K3326* variants developed both breast and ovarian cancer. Sequencing of some of her immediate family members at these loci determined that both variants were inherited from her father, who died of liver cancer and a twin sibling and paternal grandfather of the patient, both of whom died of lung cancer (a disease associated with this SNP), but whose genotypes are not available (Yufei Wang et al., 2015). A second female sibling of this patient had inherited the

ATM frameshift but not the *BRCA2* K3326* variant and developed breast cancer at the age of 46 (Figure 9B).

A.



B.

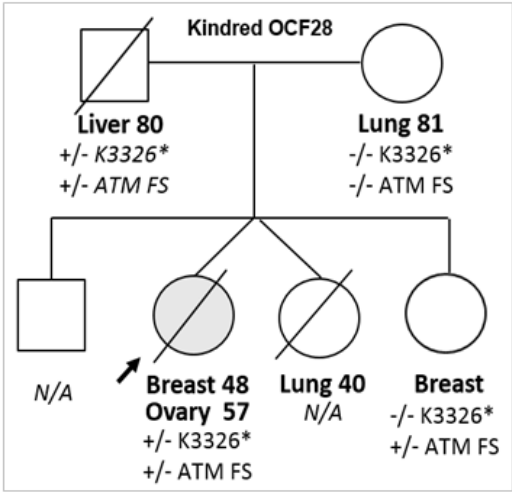


Figure 9. ***BRCA2*** **K3326* may be a modifier of penetrance A.** BRCA2 protein schematic depicting site of truncation p.K3326* variant. Variant occurs at the C-Terminus (red) which occurs in a RAD51 binding domain and site of CDK2 phosphorylation. **B.** OCF28 kindred. Arrow indicates patient OCF28-1 Kindred of proband (arrow) with p.K3326* plus pathogenic ATM frameshift shows inheritance of both alleles from an affected father.

Carriers of *BRCA2* p.K3326* and additional variants of interest

Patient ID	Gene	Consequence	Amino Acids	dbSNP ID	Variant	MAF	HGMD cancer phenotype	SIFT	PolyPhen
OCF28-1	ATM	FRAMESHIFT	CATCTG>C Exon 13	N/A	c.2503_2507del	N/A	Breast/Ovarian	N/A	N/A
	BRCA2	STOP	K3326*	rs11571833	c.9976A>T	0.009	Breast/Ovarian	N/A	N/A
	PALLD	MISSENSE	R303S	rs138897963	c.909A>T	0.001	Pancreatic	TOL	Probably Damaging
OCK1	ATM	MISSENSE	F1463C	rs138327406	c.4388T>G	0.002	Breast/Ovarian	DEL	Probably Damaging
	BRCA2	STOP	K3326*	rs11571833	c.9976A>T	0.009	Breast/Ovarian	N/A	N/A
	ERCC6	MISSENSE	M713V	rs201486862	c.2137A>G	0.00001	Basal cell carcinoma, Cockayne syndrome,	DEL	Benign
	HMMR	STOP	E352*	rs146791423	c.1054G>T	0.003	None	N/A	N/A
	RAD51D	STOP	R206*	rs387906843	c.616C>T	0.00001	Breast/Ovarian	N/A	N/A
	RECQL	MISSENSE	C321Y	rs150889040	c.962G>A	0.00001	Breast	N/A	Probably Damaging
OCN22	BRCA2	STOP	K3326*	rs11571833	c.9976A>T	0.009	Breast/Ovarian	N/A	N/A
	BUB1B	MISSENSE	E409D	rs28989188	c.1227A>C	0.0004	Gastrointestinal	TOL	Probably Damaging
OCP36	BRCA2	STOP	K3326*	rs11571833	c.9976A>T	0.009	Breast/Ovarian	N/A	N/A
	AXIN1	MISSENSE	V340M	rs143974067	c.1018G>A	0.00004	Colorectal adenoma	DEL	Probably Damaging

Table 9. Additional VUSs in carriers of BRCA2 K3326* Rare and predicted to be deleterious/damaging variants (SIFT/PolyPhen-2) found in carriers of BRCA2 p.K3326*. MAF= Minor allele frequency (ExAC, European non-Finnish) OBS= Number of times variant was observed in sample, DEL= deleterious TOL= Tolerated, N/A= Not Available. All variants listed were confirmed by Sanger DNA Sequencing.

CHAPTER 4: DISCUSSION PART I

We performed WES on 48 women with OVCA and suspected to have an inherited cancer predisposition, yet, were previously tested and found negative for known pathogenic mutations in either *BRCA1* or *BRCA2*. In doing so, we discovered pathogenic variants in *ATM* (n=2) and *FANCM* (n=1), genes currently associated with breast cancer but not OVCA, as well as in a gene recently implicated in hereditary ovarian cancer risk, *RAD51D* (n=2). These findings suggest that carriers of *ATM* and *FANCM* pathogenic mutations are possibly at elevated risk of developing OVCA as well as breast cancer and that the underling genetics of these two cancers may overlap more than previously believed. Available expression data via GTEx Portal (Broad Institute) indicate both genes have higher RPKM (reads per kilobase per million) scores in ovary tissue versus breast; *ATM* =3.6 breast, 8.7 ovary and *FANCM* =.89 breast, 1.1 ovary (*https://gtexportal.org)* indicating that these genes are expressed in ovarian cells. Furthermore, the results from WES indicate that there is clinical value of resequencing *BRCA1/2* negative individuals that fit current NCCN guidelines and whose genetic risk was assessed before the era of multi-gene panel testing.

4.1 WES highlights three likely sources of missing heritability

The majority of the high risk OVCA participants in our WES sample set did not harbor a known clinically actionable cancer predisposing mutation upon reanalysis with whole exome sequencing even in known panel genes, emphasizing the current challenge for genetic testing and counseling in clinical cancer care. Despite the large heritable component to OVCA, the majority of underlying genetic risk remains unexplained (Pharoah et al., 2013). Although many novel putative risk loci were discovered, most are rare or private familial missense mutations of

unknown clinical significance and not found in the published literature. The rarity of these variants also means that they would not be identifiable in GWAS studies.

4.1.1 Variants of Unknown Significance

Many high-risk women who undergo testing for HBOC are found to be carriers of one or more "variants of unknown significance" (VUSs), (Domchek and Weber, 2008) a rare, generally missense mutations, unannotated in their consequence to disease risk rather than a clearly pathogenic variant. Although the functional consequence of high impact variants such as nonsense and frameshift mutations are straightforward to interpret, missense mutations which result in single amino acid substitutions are of ambiguous significance. I observed suspicious missense VUSs in HBOC panel genes employing well-accepted bioinformatics techniques: *BRCA1, BRCA2, CHK2, MUTHY, MHS6, NBN, PMS2,* and most notably in *ATM* and *PALB2*. Overall, such suspicious variants in 23 of our 48 test subjects were uncovered.

The ability to assess VUSs is crucial to closing the gap in unexplained heritability while aiding in more informed clinical decisions. A common approach to implicating a VUS is by linkage analysis, whereby the causal mutation is expected to segregate with the disease in one or more families. Unfortunately, DNA samples from other affected and non-affected family members are generally not often available. A linkage analysis is also not ideal for low to moderate risk factors because these variants are not highly penetrant. Bioinformatic prediction tools for variant consequence on protein function, such as SIFT and PolyPhen, are very useful for prioritizing variants for follow up. However, *in silico* assessment tools such as these often contradict each other and are not considered to have enough sensitivity and specificity to inform clinical decisions (Richards et al., 2015). Despite the advent of detailed guidelines for variant interpretation, many variants in ClinVar list numerous testing facility submissions with

conflicting interpretations of pathogenicity. Thus, the vast majority of single nucleotide polymorphisms (SNPs) in cancer-relevant genes remain unannotated as to whether the change is deleterious to protein function and potentially disease causing.

4.1.2 Polygenic risk loci

Further complicating this issue is that under a polygenic model for hereditary cancer risk, carriers of multiple low penetrant genetic variants could be at high risk (Johnson et al., 2007), meaning much of the unexplained heritability in OVCA may be due to more than one genetic risk factor that, when inherited together, have an additive or synergistic effect. One variant in *BRCA2* (p.K3326*) stood out as a possible modifier of penetrance due to an almost five-fold increased occurrence over expected and the observation that two of the four women carrying this SNP also had a pathogenic mutation of moderate effect in a second low penetrance gene involved in DNA repair, (*ATM* and *RAD51D).* This SNP results in a 93-amino acid truncation and is reported as benign according to genetic testing services, mostly due to weak disease co-segregation in familial studies. This assessment has been questioned in recent literature due to its association with other cancers. Functional data have suggested that K3326* acts similar to wild type *BRCA2* for recombination repair and MMC sensitivity (K. Wu et al., 2005). However, the K3326* truncation is located at the C-terminus of the BRCA2 protein (exon 27), and deletion of this domain has been shown to result in reduced cellular response to stalled and collapsed replication forks, (T. M. Kim et al. 2014) hypersensitivity to gamma-radiation and premature senescence (Morimatsu, Donoho, and Hasty, 1998). Additional evidence in the literature along with our findings suggest the possibility that this variant that may be of minimal effect alone but enhances the penetrance of another moderately penetrant inherited variant in the same functional pathway. This would explain the weak genotype to phenotype correlation with this variant as

well as the observation that this variant has been found in-trans with other pathogenic *BRCA2* mutations, without causing Fanconi Anemia. Due to our small sample set, the occurrence of this SNP with additional moderate pathogenic mutations in the same pathway could be by chance. However, in agreement with the hypothesis of a role as a possible modifier of penetrance, the *BRCA2* K3326* truncation is found in The Cancer Genome Atlas (TCGA) database three times, each in OVCA patients who are all also carriers of pathogenic genetic variants; (TCGA-24-1562-01 with an *NF1* frameshift, aTCGA-13-1512-01 and TCGA-23-1026-01 with *BRCA1* frameshift mutations (*http://cancergenome.nih.gov*).

4.1.3 Novel genes yet to be implicated in risk association studies

It is likely additional risk genes exist that, when mutated, predispose to breast and/or ovarian cancer, but have yet to be implicated due to their rarity or low penetrance. In my attempt to discover novel OVCA predisposition genes, I chose to focus on genes involved in DNA repair or cell cycle control as these two dynamic and interrelated pathways are crucial to genomic stability and are the most mutated pathways in hereditary breast and ovarian cancers. In doing so, I discovered 11 high impact mutations in genes that are not featured on current HBOC risk assessment panels (*CHK1, FANCM*, *HMMR, MCM4, POLK, POLQ, RAD1, RAD52, REC8,* and *TP53I3*) but have very similar or overlapping functions to those genes on commercial panels. The finding of a pathogenic variant in *FANCM* during this specific analysis is promising as it affirms the candidate gene rationale and marks the first known case of a *FANCM* deleterious variant in an ovarian cancer patient. Of the eleven variants discovered in this analysis, five were novel. The rarity of these high impact variants is likely due to the essential natures of the DNA repair and cell cycle pathways. Knock out mouse model studies of *CHK1, MCM4*, and *RAD1* all show embryonic lethality in homozygous null mice and increased cancer incidence in

heterozygotes (Table 8), similarly to *BRCA1/2*, which makes them compelling and worthy of following up with functional studies. This study is the first of its kind to describe these germline loss of function variants in ovarian cancer patients with inherited risk. Further work should include analyzing genes in other cancer related pathways since risk loci may also occur in mechanisms not involved in DNA repair or cell cycle control.

4.2 Functional assessment is necessary to implicate novel genetic loci identified by bioinformatics tools

One key challenge facing genetic testing and counseling in clinical cancer care is the functional significance of VUSs in cancer-associated genes as well as loss of function variants in candidate risk genes. This information is necessary to provide genetics professionals with guidance for better informed patient risk evaluation, risk reduction strategies and possibly improved therapeutic modalities. While bioinformatics tools for assessment are useful for variant filtering purposes, they are not sensitive enough for clinical decisions. Ideally, missense variants predicted by bioinformatics algorithms to be 'pathogenic' or 'likely pathogenic' and novel candidate genes found with high impact mutations would be functionally tested using sensitive and specific assays that add to computational evidence for clinical insights. Because a single low-to-moderately deleterious mutation may appear inconsequential alone but could modify the penetrance of a deleterious mutation in the same pathway, combining the risk of multiple genetic variants may also lead to better risk assessment.

CHAPTER 5: RESULTS PART II

Analysis of WES data from 48 high risk women with OVCA revealed 11 loss of function variants in genes not already implicated in hereditary OVCA risk, but whose protein products are involved in DNA repair and/or cell cycle control, the two most commonly mutated pathways in HBOC. Because hereditary OVCA is rare, displays variable penetrance, and has a high degree of underlying genetic heterogeneity, the implication of a novel gene, especially of high penetrance, is unlikely to occur through case control associative studies. Therefore, I chose to functionally assess various candidate genes found mutated in our cohort. Specifically, *FANCM, CHK1, MCM4, RAD1,* and *REC8* were of interest due to their conservation and cancer-associated mouse model phenotype which includes embryonic lethality in homozygote knockouts and increased cancer incidence in heterozygotes. I was particularly interested in *TP53I3* because a very rare high impact mutation in this gene occurred twice in our cohort in unrelated individuals. Also there was very limited functional data and no mouse model phenotype in the literature. Finally, *HMMR* was also chosen because it has been shown to form a complex with *BRCA1*/*BRCA2* and common missense variations in this gene have been shown to modify the penetrance of breast cancer risk in *BRCA1* pathogenic mutation carriers (Maxwell et al., 2011).

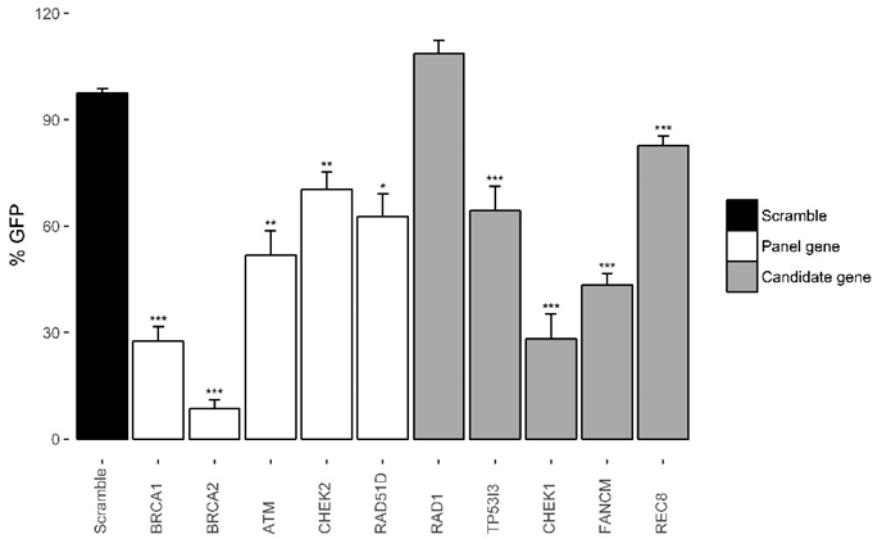
Genes chosen as positive controls for functional analyses include HBOC risk genes *BRCA1* and *BRCA2* (high risk) as well as *ATM, CHK1, RAD51D* (moderate risk). The choice of multiple positive controls with both high and moderate penetrance was to gauge the sensitivity of the assays employed for both highly and moderately penetrant genes as well as to determine whether they could distinguish between a gene of high penetrance and one of moderate penetrance. Two highly sensitive assays were employed to measure involvement in homologous

recombination repair (HRR) and sensitivity to DNA damage induction; the HRR Assay and clonogenic survival assay, respectively (see Materials and Methods sections 2.8 and 2.9). The HRR assay quantifies the efficiency of the cell's endogenous HRR pathway by inducing double strand DNA breaks (DSB) into cells with a non-functional and stably transfected GFP allele. The GFP allele will only become active once the DSB repair has occurred, therefore the amount of GFP induction proxies for the efficiency of the HRR pathway. This method accurately assays a gene's involvement in the HRR pathway by measuring the amount of GFP observed after siRNA knock down as compared to a scramble siRNA control after DSB induction. The clonogenic survival assay is a commonly applied tool to assay a gene's involvement in cell survival as well as to measure drug cytotoxicity. It is often used to determine the effectiveness of chemotherapeutic reagents under various tumor genetic profiles. All functional analyses were carried out with the use of small inhibitory RNAs (siRNA), which interfere with the expression of a targeted gene by preventing mRNA translation. Knock downs achieved by siRNA were verified by western blotting technique described in Materials and Methods section 2.7 "*Protein Detection and Quantification*." These blots are featured in the appendix of this manuscript.

5.1 Knock down of various candidate genes found mutated in cohort lead to reduced homologous recombination repair efficiency

Using the HRR protocol described in section 2.8, I compared the amount of GFP observed 48 hours after I-SCEI DSB induction among cells with siRNA knockdown of both candidate and control genes to cells with no knock down (scramble siRNA control). All conditions were adjusted for a negative control background (empty plasmid in place of I-SCEI plus scramble siRNA). Five to six replicates represent each condition to reduce inflation of Type II error. siRNA knock down of panel genes tested (*BRCA1*, *BRCA2*, *ATM*, *CHK2* and *RAD51D*),

led to a significant decrease in HRR efficiency after siRNA knock down. Notably, knock down of *BRCA1* and *BRCA2* led to the highest reduction in HRR efficiency (approximately 70% and 89% less compared to scramble control respectively (reported as mean difference MD)) while *ATM*, *CHK2* and *RAD51D* knock down led to moderate reduction in HRR efficiency (45% 27%, and 35% reduction respectively), consistent with their roles as moderately penetrant genes. siRNA depletion of candidate genes *REC8, TP53I3, CHK1* and *FANCM* all lead to a significant reduction in HRR. The greatest reduction amongst candidate genes was observed with *CHK1* which was similar to *BRCA1* and lead to a reduction of 69.2% (p=<0.001). The next largest reduction in HRR efficiency was observed with *FANCM*, (MD= 54.1, p=<0.001). Knock down of *TP53I3* led to a reduction of 33% (p=0.001) and *REC8* with a modest reduction of 14.8% (p=0.001). siRNA depletion of *RAD1* did not lead to a decrease in HRR efficiency but seemed to trend towards an increase. While this is not statistically significant, biologically it would make sense since RAD1 is believed to play a role in microhomology-mediated end joining (MMEJ) for the repair of ionizing radiation and chemicals that induce DSBs (Ma et al., 2003). Therefore, cells depleted of *RAD1* may be more reliant on the HRR pathway for DSB repair. One study has shown *RAD1* as indispensable to microhomology-mediated end joining (MMEJ) for the repair of ionizing radiation and chemicals that induce DSBs (Ma et al. 2003).

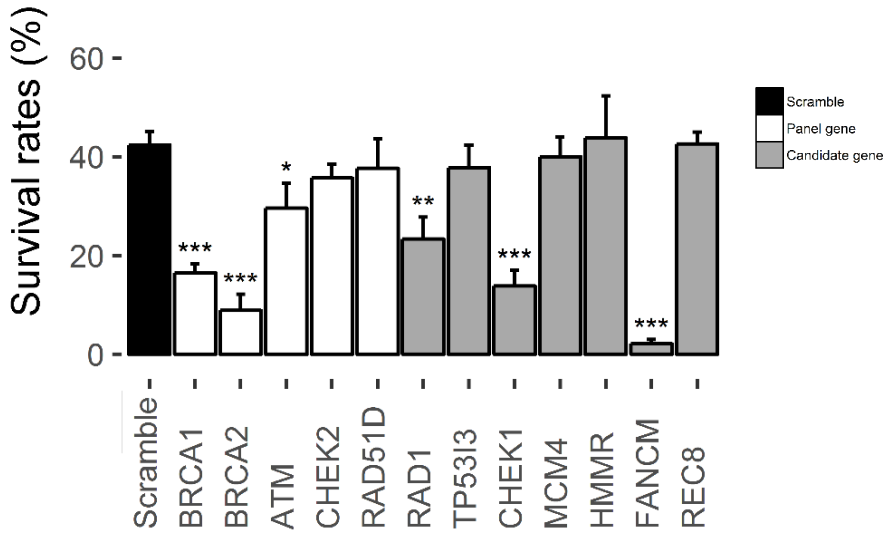


	Gene	*M*	*SD*	*MD*	*t*	*p*
	Scramble	**97.4**	**2.58**	**---**	**---**	**---**
Panel	ATM	51.9	13.43	-45.5	6.72	**0.006**
	BRCA1	27.5	8.32	-69.9	21.45	**0.000**
	BRCA2	8.6	5.02	-88.8	39.97	**0.000**
	CHK2	70.3	9.96	-27.1	5.98	**0.003**
	RAD51D	62.8	12.61	-34.7	5.452	**0.011**
Non-Panel	FANCM	43.3	6.68	-54.1	21.53	**0.000**
	CHEK1	28.2	14.04	-69.2	11.95	**0.000**
	RAD1	108.7	7.57	11.2	-2.52	0.119
	REC8	82.6	5.68	-14.8	6.40	**0.000**
	TP53I3	64.4	13.55	-33.0	7.55	**0.000**

Figure 10. HRR efficiency after siRNA knockdown. M = Mean % of GFP; SD = Standard deviation; MD = mean difference as compared to scramble control. t = statistical test for mean difference; * $p<0.05$; ** $p<0.01$; ***$p<0.001$; p-values in bold are significant. Data normalized to negative control for GFP background signal and positive control for GFP induction.

5.2 Knock down of RAD1, CHK1 or FANCM lead to a decrease in cellular viability

One key genetic mechanism in the process of tumorigenesis is the loss of heterozygosity (LOH) at tumor suppressor loci leading to the "inactivation" of genes required for the regulation of cell growth and differentiation (Ryland et al., 2015). Functional loss of genes essential for cellular viability is known to encourage LOH (Yuxun Wang et al., 2010) and is associated with increased cancer risk. For instance, ovarian breast cancer risk genes *BRCA1* or *BRCA2* are both involved in cellular viability. To identify which of the candidate genes are also involved with cellular viability, I compared the clonogenic survival of cells after siRNA knockdown of candidate and control genes to a scramble siRNA control (Figure 11). Five to six replicates represent each condition repeated to reduce inflation of Type II error. Among panel genes tested, siRNA knock down of either *ATM*, *BRCA1* or *BRCA2* lead to significant a loss of clonogenic survival. There was no loss of cellular viability observed for panel genes *CHK2* or *RAD51D*. For candidate gene *CHK1*, there was a loss of clonogenic survival similar to *BRCA1* and *BRCA2* (*CHK1* = mean survival of 13.8%, *BRCA1* = mean survival of 16.5%, *BRCA2* = mean survival of 8.8%). In addition, siRNA knock down of candidate gene *RAD1* led to a reduction in clonogenic survival that was similar to that observed with *ATM* (mean survival of 23.3% vs 29.6% respectively). Most interestingly, siRNA depletion of *FANCM* led to the greatest loss in clonogenic survival with a mean plating efficiency of just 2%.



	Gene	*M*	*SD*	*MD*	*t*	*p*
	Scramble	**42.4**	**5.44**	---	---	---
Panel	ATM	29.6	10.06	-12.8	-2.55	**0.044**
	BRCA1	16.5	3.54	-25.9	-9.78	**0.000**
	BRCA2	8.8	6.59	-33.6	-9.62	**0.000**
	CHK2	35.7	5.54	-6.7	-2.12	0.060
	RAD51D	37.7	11.94	-4.7	-0.88	0.406
Non-Panel	FANCM	2.1	1.83	-40.3	-17.2	**0.000**
	CHK1	13.8	6.37	-28.6	-8.37	**0.000**
	HMMR	43.8	17.09	1.4	0.19	0.855
	MCM4	40.0	7.87	-2.4	-0.61	0.554
	RAD1	23.3	8.82	-19.1	-4.51	**0.002**
	REC8	42.6	4.78	0.2	0.06	0.952
	TP53I3	37.8	8.98	-4.6	-1.07	0.315

Figure 11. Plating efficiency after siRNA knock down. M = Mean % survival; SD = Standard deviation; MD = mean difference as compared to scramble control. t = statistical test for mean difference; * $p<0.05$; ** $p<0.01$; ***$p<0.001$; p-values in bold are significant.

5.3 Cells deficient in CHK1, RAD1, TP53I3 or REC8 display significant sensitivity to DNA damage

Genomic stability is directly related to a cell's DNA repair efficiency. Cells deficient in DNA repair mechanisms display greater sensitivity to DNA damaging reagents. Chemotherapeutic drugs; Cisplatin, Etoposide, and Olaparib specifically exploit this vulnerability and cells deficient in tumor suppressor genes *BRCA1* or *BRCA2* are especially sensitive to these reagents. To assess whether, and to what extent, any of the candidate genes are involved in DNA repair and genomic integrity, I compared the clonogenic survival of cells with targeted siRNA depletion of candidate and panel genes to a scramble siRNA followed by exposure to Cisplatin, Etoposide, or Olaparib. These three reagents where chosen due to their different mechanisms regarding DNA damage and in relevance to current OVCA therapies. Cells were exposed to the drugs 48 hours post siRNA knock down, when the targeted depletion is greatest. Drug concentration and exposure were determined by optimizing for an IC_{50}, which for the purposes of these experiments, is the concentration required for 50% cytotoxicity in a scramble siRNA knock down control. The IC_{50} treatment exposure for each reagent was as follows: Cisplatin= 2μM for 2 hours, Etoposide= 10μM for 4 hours, and Olaparib 10μM for 4 hours. The clonogenic survival after drug exposure for each condition was adjusted to the plating efficiency observed with the same siRNA knock-down and without drug exposure. Five to six replicates represent each condition repeat to reduce inflation of Type II error. The assessment of candidate gene *FANCM* to these reagents was not possible due to the extreme loss of cellular viability that occurred after siRNA depletion.

5.3.1 Cells deficient in CHK1, RAD1 or TP53I3 display increased sensitivity to Cisplatin

Cisplatin generates interstrand cross links (ICLs) which are covalent bonds between both strands of a DNA duplex. ICLs inhibit crucial processes such as DNA replication, ultimately leading to chromosomal instability. ICLs are repaired by HRR and thus, cells deficient in this pathway are highly sensitive to DNA-damaging agents such as Cisplatin. As expected, all panel genes included in this assay *(BRCA1, BRCA2, ATM, CHK2,* and *RAD51D*) were demonstrated to be involved in HRR and ICL repair as targeted siRNA against their respective mRNA products led to significantly increased cytotoxicity with Cisplatin exposure (Figure 12A, Table 10). Specifically, loss of BRCA1 or BRCA2 leads to the most sensitivity, each with a mean difference of approximately 40% increased cytotoxicity as compared to the scramble control. This observation is consistent with their status as highly penetrant in cancer susceptibility when mutated. Cells with knock down of candidate genes: *CHK1*, *RAD1*, or *TP53I3* also displayed significantly increased sensitivity to cisplatin exposure equal of greater to that observed in the moderately penetrant panel genes *ATM, RAD51D* and *CHK2* (Figure 12A, Table 10). Loss of *RAD1* led to the most sensitivity observed amongst candidate genes with a mean viability of 18.3% which is a 33% increased cytotoxicity as compared to the scramble control ($p=<0.001$). Depletion of *CHK1* and *TP53I3* both led to a mean increase in cisplatin cytotoxicity of approximately 27% ($p=<0.001$, $p=<0.001$ respectively.) There was no increased cisplatin cytotoxicity observed after knock-down of *REC8*, *HMMR* or *MCM4*.

5.3.2 Cells deficient in CHK1, REC8 or RAD1 display increased sensitivity Etoposide

Etoposide is a topoisomerase II (topoII) alpha inhibitor approved for clinical use as a chemotherapeutic reagent in platinum resistant OVCA. Topoisomerase II enzymes are responsible for simultaneously cleaving both stands of the DNA double helix for the

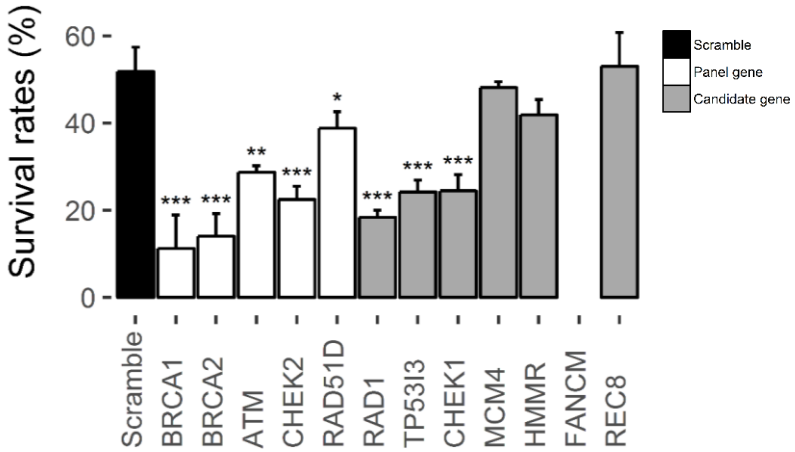
management of entangled and supercoiled DNA. These enzymes are essential for DNA replication. Inhibition of topoII by Etoposide prevents the re-ligation of cleaved DNA and therefore leads to DNA DSBs. Etoposide also results in single-strand DNA breaks (SSBs), since it prevents the re-ligation of the stands independently of each other (Yang et al. 2009). *BRCA1* and *BRCA2* deficient cells have been described as sensitive to etoposide treatment (Treszezamsky et al., 2007). Consistent with these reports, both *BRCA1* and *BRCA2* deficient cells displayed significant sensitivity to Etoposide in this assay. Knock down of *CHK2* or *RAD51D* also led to increased Etoposide sensitivity (with approximately 30% and 14% increased cytotoxicity respectively). Among non-panel candidate genes, *CHK1, RAD1* and *REC8* knock down all resulted in increased cytotoxicity in response to Etoposide. *CHK1* deficient cells displayed a mean survival of 34.9% ($p=0.024$), *RAD1* 24.5% ($p=<0.000$) and REC8 35.4% ($p=0.018$), which is an increase of approximately 14%, 24% and 13% in cytotoxicity versus the scramble control respectively. (Figure 12B)

5.3.3 Cytotoxicity to Olaparib is specific to BRCA1 and BRCA2 deficient cells.

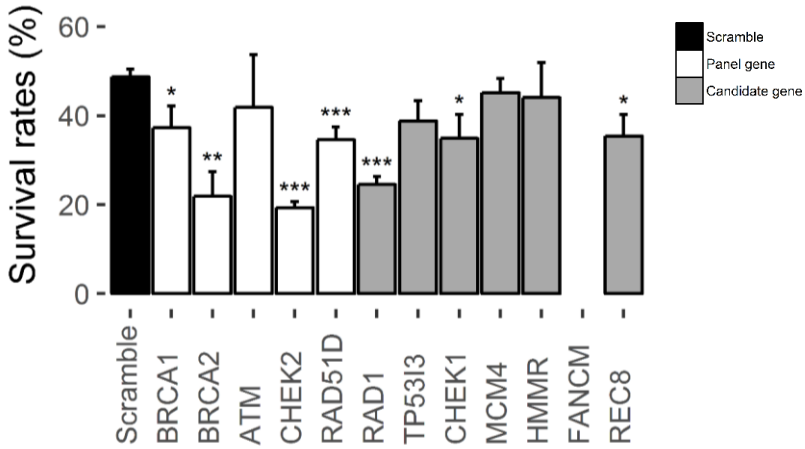
Olaparib is a PolyADP-ribose polymerase (PARP) inhibitor approved for clinical as an adjunct to platinum-based therapies in patients with *BRCA1* or *BRCA2* deficiencies. Tumor cells with mutant *BRCA1* and/or *BRCA2* have demonstrated to be up to 1000 times more sensitive to PARP inhibitors as compared to WT cells (Farmer et al. 2005; Bryant et al. 2005). Cancer cells deficient in HRR are more dependent on the PARP proteins which are involved in ssDNA break repair by Base Excision Repair (BER). Inhibiting this pathway leads to synthetic lethality as the cells loses its back up mechanism to repair DNA. In clinical trials, the use of PARP inhibitors have shown to improve the progression-free survival when added to the treatment of women with breast or ovarian cancer responsive to platinum, which induces DSBs (Robson et al., 2017;

Ledermann, 2016; Evans and Matulonis, 2017). Olaparib used as a monotherapy has shown to be effective in patients with germline *BRCA1/2* mutation and advanced cancer (G. Kim et al. 2015; Kaufman et al., 2015). In this experiment, cells were treated with Olaparib in the absence of cisplatin and increased cytotoxicity was specific to BRCA1 and BRCA2 deficient cells (Figure 12C, Table 10). *ATM, CHK2* and *RAD51D* panel genes have demonstrated a less prominent role in HRR as compared to *BRCA1* or *BRCA2* (Figure 10). This observation may indicate that sensitivity to Olaparib requires a heavier reliance on the BER pathway. It is also possible that deficiency of *ATM, CHK2* or *RAD51D* would lead to increased Olaparib sensitivity, but only in conjunction with platinum-based therapy.

A. *Cisplatin*



B. *Etoposide*



C. *Olaparib*

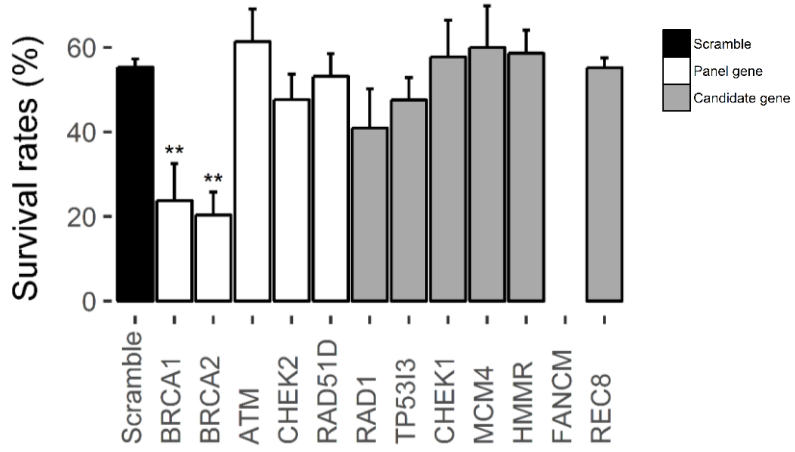


Figure 12. Clonogenic survival rates after drug exposure by siRNA knock down A. Cell Survival after exposure to Cisplatin by siRNA knockdown. All values adjusted to mean plating efficiency after siRNA knockdown alone. **B.** Cell Survival after exposure to Etoposide by siRNA knockdown. All values adjusted to mean plating efficiency after siRNA knockdown alone. **C.** Cell Survival after exposure to Olaparib by siRNA knockdown. All values adjusted to mean plating efficiency after siRNA knockdown alone. * $p<0.05$; ** $p<0.01$; ***$p<0.001$.

		Condition (Survival rates)														
		Cisplatin					Etoposide					Olaparib				
	Gene	*M*	*SD*	*MD*	*t*	*p*	*M*	*SD*	*MD*	*t*	*p*	*M*	*SD*	*MD*	*t*	*p*
	Scramble	**51.8**	**11.19**	---	---	---	**48.8**	**3.30**	---	---	---	**55.3**	**3.94**	---	---	---
Panel	ATM	28.7	2.93	-23.1	-4.74	**0.003**	41.9	23.77	-6.9	-0.70	0.511	61.4	15.42	6.0	0.77	0.495
	BRCA1	11.2	15.38	-40.6	-5.23	**0.001**	37.3	9.71	-11.5	-2.74	**0.033**	23.7	17.54	-31.6	-4.31	**0.006**
	BRCA2	14.0	10.44	-37.8	-6.05	**0.000**	21.9	11.13	-26.9	-5.21	**0.004**	20.4	10.85	-35.0	-6.84	**0.001**
	CHEK2	22.4	6.12	-29.4	-5.52	**0.001**	19.3	2.66	-29.4	-16.37	**0.000**	47.7	12.10	-7.7	-1.36	0.235
	RAD51D	38.8	7.61	-13.0	-2.36	**0.043**	34.6	5.77	-14.2	-5.22	**0.001**	53.2	10.68	-2.2	-0.47	0.656
Non-Panel	FANCM	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
	CHEK1	24.4	7.38	-27.4	-5.00	**0.001**	34.9	10.83	-13.8	-3.00	**0.024**	57.7	17.46	2.4	0.33	0.754
	HMMR	41.9	6.99	-10.0	-1.85	0.100	44.1	15.84	-4.7	-0.71	0.508	58.7	10.88	3.3	0.70	0.507
	MCM4	48.1	2.71	-3.7	-0.78	0.465	45.1	6.40	-3.6	-1.23	0.256	60.0	19.67	4.7	0.57	0.592
	RAD1	18.3	3.31	-33.5	-7.03	**0.000**	24.5	3.55	-24.2	-12.25	**0.000**	41.0	18.53	-14.4	-1.86	0.117
	REC8	53.0	15.54	1.2	0.14	0.891	35.4	9.64	-13.3	-3.21	**0.018**	55.2	4.70	-0.2	-0.05	0.963
	TP53I3	24.2	5.44	-27.6	-5.44	**0.001**	38.8	9.11	-10.0	-2.32	0.069	47.6	10.51	-7.7	-1.69	0.139

Table 10. Statistical data for clonogenic survival rates after Cisplatin, Etoposide and Olaparib treatment by siRNA knockdown. M = Mean % survival; SD = Standard deviation; MD = mean difference as compared to scramble control. t = statistical test for mean difference; * $p<0.05$; ** $p<0.01$; ***$p<0.001$; p-values in bold are significant.

5.4. Estimating penetrance through a 2-dimensional model of cell viability and DNA sensitivity

Most cancer-associated risk loci have incomplete penetrance. The penetrance of a pathogenic variant is determined by its associated lifetime risk with disease. Genes are often described as having low, moderate, or high penetrance based on the lifetime risk associated with a loss of function variant. *BRCA1* is the most highly penetrant gene in hereditary breast and ovarian cancer syndrome with an 80% lifetime risk of breast cancer and 50% risk of OVCA for LOF variants. Accurate estimates of genetic penetrance are necessary to inform clinical decisions. Therefore, it would be optimal to evaluate novel risk genes in a manner that would provide an indication of penetrance employing a functional test.

The two genes with the highest known penetrance for HBOC in the literature, *BRCA1* and *BRCA2*, are important to both cellular viability and DNA repair. Loss of function of either leads to a profound reduction of cellular viability, and increased sensitivity to DNA damaging reagents as observed in the various functional assays employed in this study. Functional depletion of genes that are moderate in their penetrance, such as *ATM*, *RAD51D* and *CHK2*, lead to more moderately increased DNA damage sensitivity as compared to *BRCA1/BRCA2*, and may or may not impact cell viability in the absence of a cytotoxic reagent. Because cell viability and DNA damage sensitivity can be observed independently of each other yet are both phenotypes of a *BRCA-like* tumor suppressor gene, plotting genes using a two-dimensional graph based on these phenotypical outcomes may lead to a separate geographical clustering of high and moderately penetrant genes in HBOC risk. By plotting genes with known penetrance, we can assess the potential of this method. If accurate, we can then use this scale to estimate the likely risk/penetrance of candidate risk genes that act as a tumor suppressor in a *BRCA-like* manner.

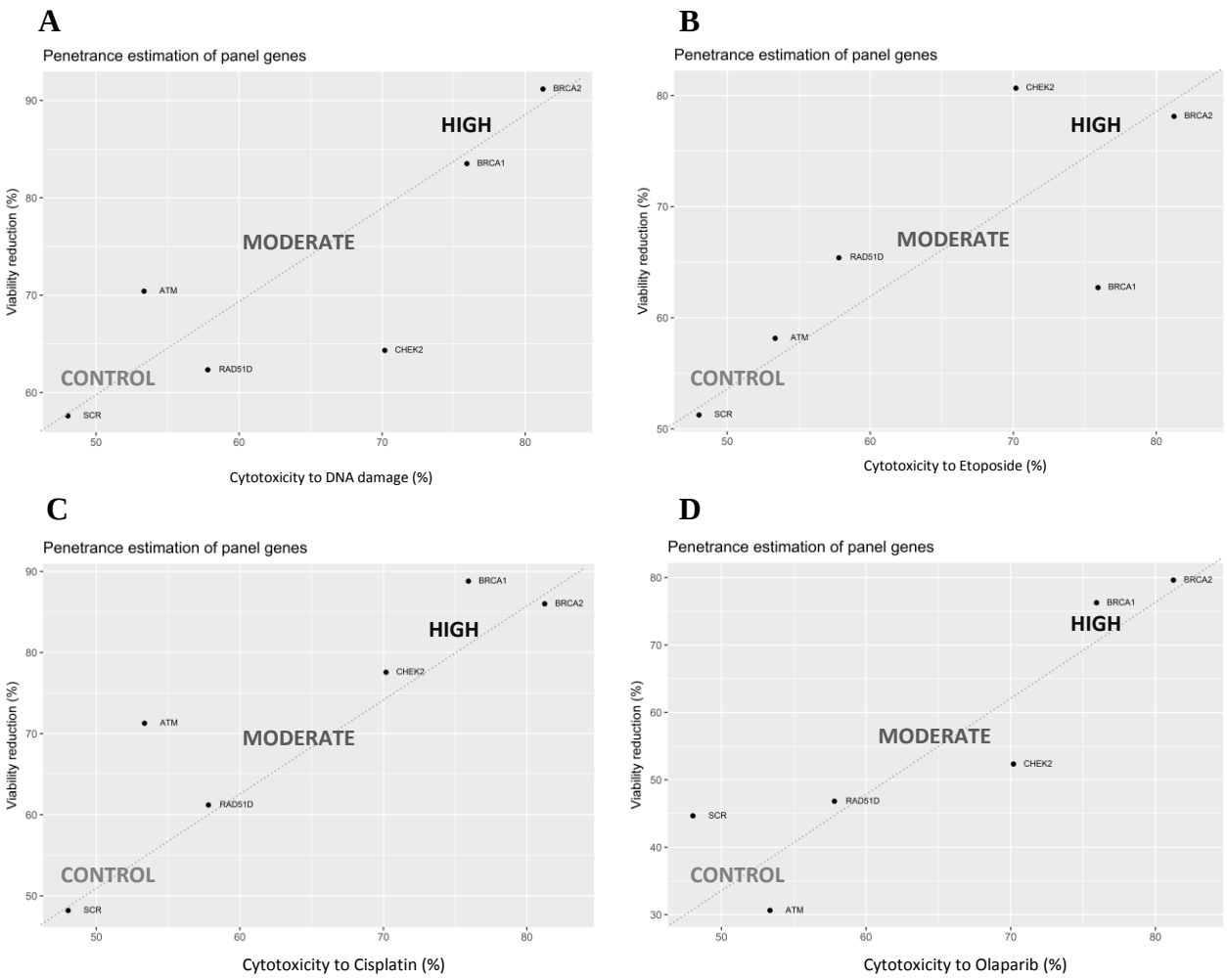
5.4.1 Plotting panel genes based on cell viability and DNA damage sensitivity after siRNA knock down correctly differentiates between moderate and high penetrance

The three different chemotherapeutic drugs (Cisplatin, Etoposide and Olaparib) employed for the clonogenic survival assays each work via three different mechanisms to exploit repair deficiencies (see section 1.11 Targeted therapy in clinical cancer care). *BRCA1* and *BRCA2* are the most highly penetrance HBOC associated genes and *BRCA1* and *BRCA2* deficient cells are particularly sensitive to all three of these chemotherapeutic reagents. Cells deficient in tumor suppressor genes that are *BRCA*-like but moderate in penetrance, such as *ATM*, *RAD51D* and *CHEK2*, display significant but less or no increased cytotoxicity to each of these three reagents. Therefore, in assessing the extent to which a gene is *BRCA*-like, it seems rationale to take into consideration the sensitivity to all three of these reagents; Cisplatin, Etoposide and Olaparib versus a single reagent alone.

To test this assumption, four plots were generated each with reduction in cell viability observed without drug exposure on the Y axis and cytotoxicity observed after DNA damage on the X axis. Figure 13A features a plot that includes the pooled cytotoxicity data across all three tested reagents; Cisplatin, Etoposide and Olaparib. The subsequent plots were generated by only accounting for cytotoxicity to each drug alone; Cisplatin; Figure 13B, Etoposide; Figure 13C, and Olaparib; Figure 13D. Since we know the penetrance of the panel genes plotted (listed in Figure 13), we can see that the first plot (Figure 13A) which features the pooled sensitivity data across all three reagents is most accurate. With this plot, highly penetrant genes *BRCA1* and *BRCA2* fall in top most section of both DNA damage sensitivity and reduction of cellular viability (upper right), while the scramble control falls in the lowest section of both conditions

(lower left). Panel genes *ATM*, *RAD51D* and *CHK2*, known to be of moderate penetrance, fall in the middle of the graphical representation.

The statistical data generated from pooling all conditions (cell viability after knockdown alone, and cytotoxicity to Cisplatin, Etoposide and Olaparib after siRNA knock down) is represented in Table 11. The overall survival of cells with siRNA scramble control was 49.6%, which is expected due to the IC_{50} optimization of drugs and normal plating efficiency of HeLa cells. The survival of *BRCA1* deficient cells across all conditions was 22% (p=<0.00) and *BRCA2* was 15% (p=<0.00). All moderately penetrant genes fell in a range of 30-39% survival across all conditions and were all statistically significant. The statistical data and geographical clustering based on these data observed across panel genes tested are concordant with the current penetrance estimates in the literature, and with what clinicians refer to for genetic counseling and risk management. Therefore, this methodology could potentially be useful in estimating candidate gene penetrance.



	Breast	Ovarian	Penetrance
ATM	17%-52%	unknown	Moderate
CHK2	23-48%	unknown	Moderate
RAD51D	unknown	14.8%	Moderate
BRCA1	46%-87%	39%-63%	High
BRCA2	43%-84%	15%-27%	High

Lifetime risks for breast and ovarian cancer determine penetrance

Stats provided by Myriad Genetics (https://myriad.com/)

Figure 13. Plotting panel genes based on cell viability and pooled DNA damage sensitivity after siRNA knock down correctly differentiates between moderate and high penetrance. **A.** Plot with pooled DNA damage sensitivity data across all three reagents correctly indicates *ATM*, *RAD51D*, *CHK2* as moderately penetrance and *BRCA1* and *BRCA2* as high penetrant. **B.** Etoposide incorrectly estimates *CHK2* as highly penetrant and *BRCA1* as moderately penetrant. **C.** Cisplatin incorrectly estimates *CHK2* as a highly penetrant gene **D**. Analysis of Olaparib alone groups *ATM* with scramble control.

5.4.2 *CHK1* and *RAD1* predicted as moderate to high and *TP53I3* as moderate penetrant genes

As previously discussed, plotting panel genes based on cell viability and pooled cytotoxicity to DNA damage after siRNA knock down may differentiate between moderate and high penetrance. In total, six control data points; scramble (SCR), *ATM*, *CHK2*, *RAD51D*, *BRCA1* and *BRCA2* represent an accurate comparison for estimating the penetrance of candidate genes. I therefore plotted these data points once more alongside non-panel candidate genes. Again, with reduction in cell viability observed without drug exposure on the Y axis and cytotoxicity observed after DNA damage (Cisplatin, Etoposide, Olaparib) on the X axis (Figure 14). Table 11 lists the statistical data from overall survival across all conditions by siRNA knockdown.

In this analysis, *MCM4*, *HMMR* and *REC8* were observed to cluster with the scramble (SCR) control suggesting they are unlikely to be *BRCA*-like tumor suppressor genes. There is no statistical difference between these candidate genes and the scramble control in overall mean survival. While *REC8* deficient cells showed low but significant sensitivity to Etoposide treatment, the difference between scramble control and pooled DNA treatment sensitivity plus plating efficiency (cell viability after knock-down) was insignificant ($p=0.06$). In the plot for estimating penetrance, candidate gene *TP53I3* clustered with moderately penetrant genes and had a survival of 36% ($p=<0.001$) across all conditions versus the survival of cells with siRNA scramble control at 49.6%. *CHK1* fell in a graphical position indicative of a gene with moderate to high penetrance as the loss of cellular viability without DNA damage was observed as similar to *BRCA1* and *BRCA2*, however the sensitivity observed in response to DNA damaging reagents was similar to moderately penetrant genes. The mean pooled survival for cells deficient in *CHK1*

for all conditions was 32.7% (p=0.001). Of all candidate genes assayed, *RAD1* was shown to be the most impactful regarding cell survival and cytotoxicity to DNA damage. Cells deficient for this gene had a survival rate of 26.8% (p=<0.001) across all conditions and this gene fell close to *BRCA1* and *BRCA2* on the graphical representation of estimated penetrance, indicating that it may be a highly penetrant gene (Table 11).

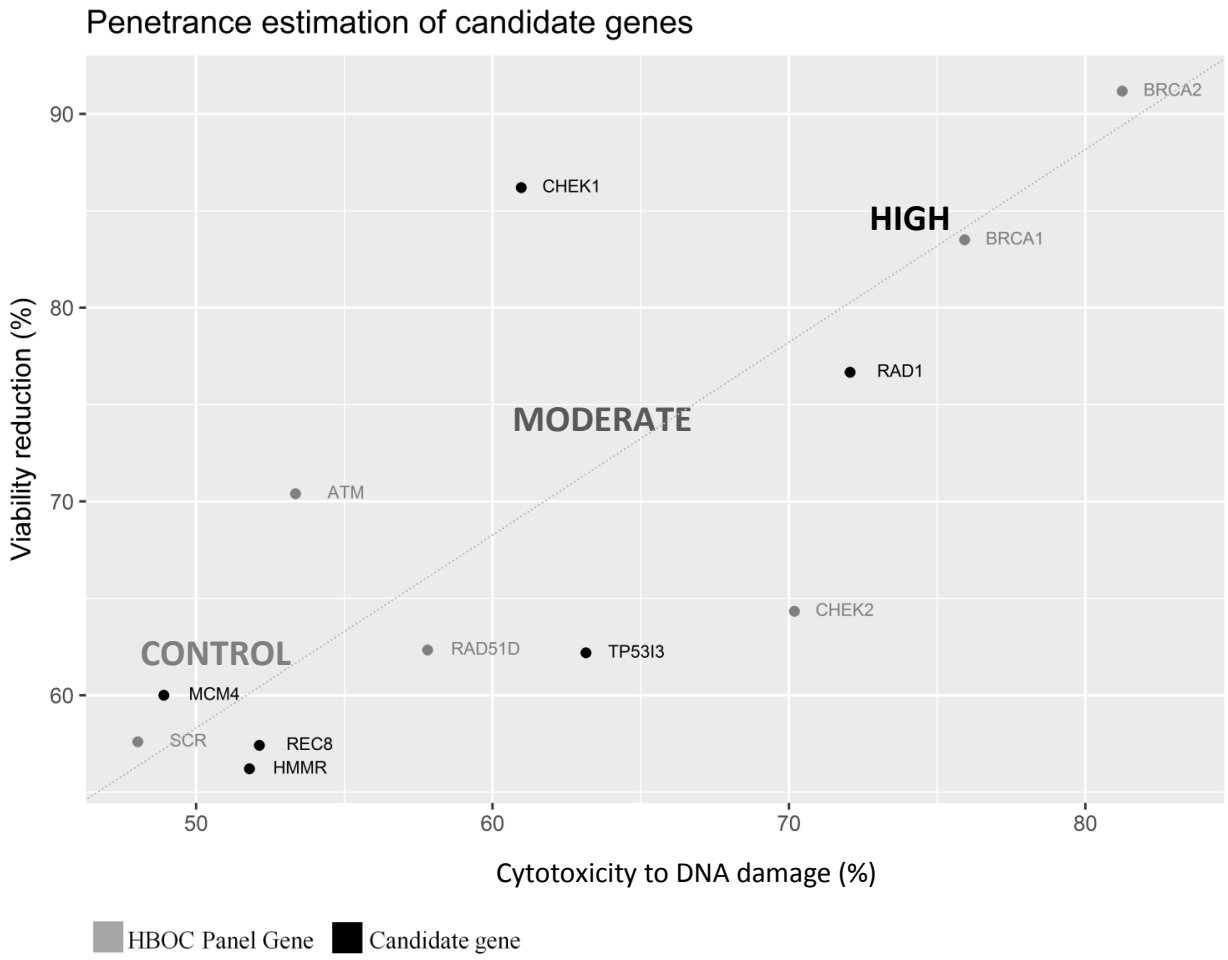


Figure 14. Penetrance estimates based on clustering with known risk genes

Pooled survival rates (%)

all conditions

	Gene	*M*	*SD*	*MD*	*t*	*p*
	Scramble	**49.6**	**8.3**	**---**	**---**	**---**
Panel	ATM	37.5	19.89	12.0	-2.60	**0.015**
	BRCA1	22.2	15.54	27.4	-7.71	**0.000**
	BRCA2	14.8	10.27	34.7	-13.34	**0.000**
	CHK2	30.1	12.95	19.5	-6.30	**0.000**
	RAD51D	38.6	12.92	11.0	-3.72	**0.001**
Non-Panel	FANCM	2.1	1.83	47.5	-26.9	**0.000**
	CHK1	32.7	19.70	16.8	-3.89	**0.001**
	HMMR	46.5	13.79	3.0	-0.96	0.342
	MCM4	48.4	12.17	1.2	-0.41	0.684
	RAD1	26.8	13.12	22.8	-7.31	**0.000**
	REC8	44.1	11.51	5.5	-1.89	0.066
	TP53I3	36.0	11.61	13.6	-4.90	**0.000**

Table 11. Statistical data for pooled survival rates across all clonogenic survival conditions; no drug, Cisplatin, Etoposide and Olaparib. M = Mean % survival, SD = Standard deviation, MD = mean difference as compared to scramble control. *t* = statistical test for mean difference, *p*-values in bold are significant.

CHAPTER 6: DISCUSSION PART II

6.1 Functional analyses compliment bioinformatics and strengthen cases for various novel risk genes

During WES analysis, various candidate genes were found with loss of function mutations in women with OVCA and high risk of genetic inheritance. While compelling variants are often identified through WES/WGS, they remain putative risk loci until proven otherwise. Because high risk variants are rare, and hereditary OVCA has a great deal of heterogeneity, implicating a novel gene or variant based on enrichment in cases versus controls is unlikely. Additionally, implicating a variant based on segregation is not ideal due to incomplete penetrance and lack of informative family members. Bioinformatic tools alone are not sensitive enough to direct clinical decisions but are useful for variant filtering purposes.

This study is unique in its ability to identify novel risk loci for two main reasons; first, WES was carried out on a very select group of patients with high risk of genetic inheritance, yet with no known pathogenic variant. Second, candidate loci identified by well-established bioinformatics techniques were followed up by functional assessment using sensitive wet lab techniques. This approach identified four novel genes; *FANCM, CHK1, RAD1* and *TP53I3* as having the *BRCA*-like phenotype typically observed in tumor suppressor genes commonly mutated the germline of women with inherited risk of breast and/or ovarian cancer.

In this study, siRNA knock-down of *FANCM* led to a reduction in homologous recombination repair and large loss of clonogenic survival similar to that observed in *BRCA1* or *BRCA2* deficient cells. Because of the large loss in cell survival after knock-down, it could not be assayed for cytotoxicity to Cisplatin, Olaparib, or Etoposide after knock-down and therefore an estimate of this gene's penetrance could not be established. However, the functional data

from this study, along with the family pedigree of the *FANCM* carrier (Figure 7), and association with familial breast cancer in recent literature (Peterlongo et al. 2015) is supportive of high risk. Additionally, mouse model data describes *FANCM* homozygous knock outs as having decreased life span and increased cancer incidence (Bakker et al., 2009).

Candidate gene *CHK1,* like panel gene *CHK2,* encodes for a serine/threonine protein kinase required for checkpoint-mediated cell cycle arrest and activation of DNA repair by homologous recombination repair. This gene is highly conserved, and the c.1564-1565insA frameshift variant identified by WES in this study sample is novel. In this study, siRNA knock down of *CHK1* lead to a loss of HRR efficiency, reduction in cellular viability, and increased sensitivity to Cisplatin and Etoposide similar to that of observed in *BRCA1/BRCA2* deficient cells. Loss of function of this gene is predicted to be moderate to highly penetrant in cancer risk as it clusters with known high and moderate penetrant tumor suppressor genes on a 2-dimensional model of cell viability and cytotoxicity to DNA damage. These results are consistent with mouse model phenotypes describing homozygote nulls are embryonic lethal, and heterozygote knock-outs display enhanced tumorigenesis (Q. Liu et al., 2000). Recently, *CHK1* has been identified as an important biomarker for chemotherapy response in breast cancer (Al-Kaabi et al. 2015) and the deletion of *CHK1* is a common genetic event that occurs in the beginning stages of breast cancer development (Mu et al., 2011). Loss of heterozygosity is likely to have occurred in the carrier of the *CHK1* frameshift. Unfortunately, the patient declined tumor tissue usage, and this could not be confirmed. The fact that *CHK1* has not already been implicated in in hereditary risk of ovarian or breast cancer is likely because pathogenic variants in this gene are extremely rare.

Another novel loss of function variant was uncovered in *RAD1*. Knock out mouse models of this gene describe embryonic lethality in homozygotes and larger, more numerous, earlier onset skin tumors with DMBA-TPA treatment in heterozygotes (L. Han et al., 2010). In this study, siRNA knock down of *RAD1* led to decreased cellular viability and increased sensitivity to Cisplatin, and Etoposide similar to that observed with knock down of *BRCA1* and *BRCA2*. However, *RAD1* depletion did not lead to a reduction in HRR efficiency. This indicates that *RAD1* may be involved in DNA repair via a mechanism other than HRR. Most of what is known regarding RAD1 functional activity is derived from yeast studies. RAD1 is a part of the 9-1-1 cell cycle checkpoint complex to arrest cellular proliferation in the presence of incomplete DNA replication or damaged DNA. This complex has also been shown to participate in DNA repair by forming a clamp to facilitate resection of DNA double strand breaks points (Ngo and Lydall 2015). One study has shown RAD1 as indispensable to microhomology-mediated end joining (MMEJ) for the repair of ionizing radiation and chemicals that induce DSBs (Ma et al. 2003). Another study in yeast identified *RAD1* mutants as hyper sensitive to platinum (Perego et al. 1998). While the human orthologue of yeast Rad1 is not well described in the literature, the observation of a germline *RAD1* LOF variant in an OVCA patient considered at high risk of genetic inheritance, plus the results of this study's functional analyses in a human cervical cell line makes a compelling case for this gene as a risk factor. The scale of penetrance developed in this study estimates the LOF of *RAD1* to be high risk. The carrier of the *RAD1* frameshift variant developed OVCA at the age of 65, and had a family history of colon, breast (n=2), prostate, lung, and leukemia all of which on one parental side of the family, which is indicative of a highly penetrant germline risk variant. Unfortunately, segregation analysis was not possible since DNA samples from her deceased family members are not available.

TP53I3 is an oxidoreductase-like protein and an inducer of ROS, that is transcriptionally activated by the tumor suppressor TP53 and likely to be involved in TP53-mediated apoptosis (Zhang et al. 2015). A nonsense SNP in this gene was observed twice in this study sample amongst unrelated individuals, despite its rarity (rs145078765, p.S252*, MAF= 0.0009). The functional analysis of this gene during this study indicated that loss of *TP53I3* leads to a moderate reduction in HRR efficiency as well as increased cell sensitivity to Cisplatin. Loss of function of this gene is predicted to be moderately penetrant in cancer risk as it clusters with known moderately penetrant tumor suppressor genes on a 2-dimensional graphical representation of cell viability and DNA sensitivity.

Another candidate gene assessed in this study was *REC8*, which encodes for a cohesin complex protein required for the structural maintenance of chromosomes during meiosis. Cohesions are necessary to join sister chromatids together until DNA replication is complete. Rec8 functional studies in yeast described it as a meiosis-specific (Yoon et al. 2016). However, in this study, siRNA reduction of REC8 led to a decrease in DNA repair by homologous recombination and increased sensitivity to Etoposide. Because the pooled data across all clonogenic survival conditions did not reach not significance (p=.06), this may indicate that *REC8* LOF would be of minimal risk, and possibly a modifier of penetrance. However, as mentioned previously, the methods employed in this study are specific to capture a *BRCA*-like tumor suppressor phenotype. It is possible that the tumor suppressor functions of REC8 are mostly independent of DNA repair pathway. Recently, investigators identified *REC8* as a tumor suppressor gene epigenetically downregulated in gastric cancer (Yu et al. 2017). Another study showed that epigenetic silencing of *REC8* was robustly associated with PI3K pathway alterations

in thyroid tumors, possibly encouraging the oncogenic properties of the PI3K/AKT/mTOR pathway, which is important to cell cycle regulation (D. Liu et al., 2015).

6.2 A Novel scale for estimating penetrance through a 2-dimensional model of cell viability and DNA sensitivity

Most cancer-associated risk loci have incomplete penetrance. The penetrance of a pathogenic variant is determined by its associated lifetime risk with disease, and accurate estimates of genetic penetrance are necessary to inform clinical decisions. Genes are often described as having low, moderate, or high penetrance based on the lifetime risk associated with a loss of function variant. Unfortunately, lifetime risk estimates are difficult to obtain with rare variants and even well-established cancer risk loci can have unknown penetrance. Currently, there exists no laboratory-based functional method for the specific purpose of estimating penetrance of genetic loci in HBOC.

Candidate tumor suppressors are often assayed in conjunction with a *BRCA1* or *BRCA2* positive control since these genes are the two most highly penetrant in HBOC risk. The term "*BRCAness*" or *"BRCA-like"* has come to describe other tumor suppressor genes that lead to similar phenotypes with loss of function, such as sensitivity to DNA damage and loss of cell viability. Various HBOC risk loci that possess *BRCA-like* tumor suppressor properties have been implicated in the last decade. Many risk loci are described as moderately penetrant due to lifetime risks estimated to be much higher than the general population, yet much lower than *BRCA1* or *BRCA2*. In this study, all lab experimental designs incorporated the use of five positive controls for the comparison of novel candidate risk genes to two highly and three moderately penetrant established risk genes in the context of *BRCA-like* tumor suppressor properties.

The assays employed in this study were sensitive enough to implicate all five established HBOC cancer risk genes tested. Importantly, a pooled analysis of cell viability and cytotoxicity to DNA damage was able to clearly distinguish between moderate and high-risk genes in a 2-dimentional geographical representation. Notably, this technique estimated candidate genes *CHK1* and *RAD1* as moderate to high and *TP53I3* as moderately penetrant genes.

While preliminary, this scale has the potential to functionally validate additional candidate genes identified by other WES studies and possibly estimate the penetrance of certain established HBOC risk genes with unknown lifetime risk profiles (*RAD50, BRIP1, BARD1,* etc.) This approach could eventually serve as a tool to better inform clinical management of certain patients. Genetic testing clinics that engage in research may choose to employ this methodology to help in closing the gap in missing heritability of this disease.

6.3 Limitations and Future Directions

It is important to note the various limitations of this study. Firstly, because WES targets exonic DNA, it was only possible to analyze the protein coding and closely surrounding regions on the genome in these high-risk women. While bioinformatic analysis was performed on the 1000 bps captured upstream of the exons of panel genes, no suspicious variants were uncovered in either promoter or splice donor/acceptor loci. Additionally, large copy number variants (CNVs) are not readily detectable through WES data and could be a source of missing heritability. Using ExomeDepth, an R package which relies on read depth to indirectly infer deletions or duplications, I was able to analyze all exomes for possible CNVs in panel genes, but no true calls were detected (data not shown).

Additionally, the candidate gene analysis performed was primarily focused on DNA repair and cell cycle control as these are the two most commonly mutated pathways in HBOC. It

is likely that risk loci occur outside of these pathways and further analyses of these exomes should be performed with this in mind. However, one important consideration in pathway analysis of WES data is that it is far easier to determine that a variant leading to a loss of function versus a gain of function by sequence change.

Furthermore, it must be emphasized that the penetrance estimates derived from functional studies are specific to a loss of function risk loci in *BRCA-like* tumor suppressor genes involved in cell viability and/or DNA repair. Accordingly, the functional analysis employed in this study suggests that candidate genes *REC8, HMMR,* and *MCM4* lack the "*BRCA-ness*" tumor suppressor phenotype typically observed in breast and ovarian cancer risk genes. However, we cannot say for certain that they are not risk factors by other means. Still, the overall approach may be applicable to develop similar penetrance estimate scales specific to other cancer related pathways. Additionally, the use of siRNA for gene depletion is not optimal since the various siRNA have different knock down efficiencies. This makes it difficult to accurately compare phenotypic outcomes by gene knockdown. Finally, because this methodology is new, it should be further validated, refined, and replicated in additional cell lines, and preferably by employing gene knock out techniques in place of mRNA depletion.

APPENDIX

Gene	Amino Acids	dbSNP ID	MAF	OBS	HGMD (cancer phenotype associated with gene)	SIFT	Polyphen
APEX1*	P248L	rs201100630	0.00005	1	Head and Neck	DEL	Benign
AXIN1	V340M	rs143974067	0.00004	1	Colorectal adenoma	DEL	Probably Damaging
BUB1B	E409D	rs28989188	0.0004	1	Gastrointestinal	TOL	Probably Damaging
CASP10*	I406L	rs80358239	0.004	1	Autoimmune lymphoproliferative syndrome II	TOL	Possibly Damaging
ERCC4*	E875G	rs1800124	0.019	1	Lung, Cockayne, Xeroderma pigmentosa, Breast/Ovarian, Fanconi anaemia,	DEL	Possibly Damaging
ERCC6	M713V	rs201486862	0.00006	1	Cockayne syndrome, basal cell carcinoma,	DEL	Benign
EXO1*	D270V	rs201509012	0.0005	1	Colorectal	DEL	Possibly Damaging
EXO1*	G759E	rs4150001	0.009	1	Colorectal	TOL	Benign
FANCA*	T475M	N/A	N/A	1	Fanconi Anemia	DEL	Possibly Damaging
FANCA*	A602G	N/A	N/A	1	Fanconi Anemia	DEL	Possibly Damaging
FANCF*	P320L	rs45451294	0.017	2	Fanconi Anemia	TOL	Probably Damaging
MLH3*	V741F	rs28756990	0.006	1	Colorectal, Breast/Ovarian,	TOL	Possibly Damaging
PALLD*	R303S	rs138897963	0.001	1	Pancreatic	TOL	Probably Damaging
PMS1*	T75I	rs61756360	0.0008	1	Breast/ovarian	DEL	Probably Damaging
RAD50*	T191I	rs2230017	0.0007	1	Breast/Ovarian	DEL	Benign
RBL1*	R199H	N/A	N/A	1	Multiple adenoma	DEL	Probably Damaging
RBL1*	E624Q	N/A	N/A	1	Multiple adenoma	TOL	Possibly Damaging
RECQL	C321Y	rs150889040	0.00001	1	Breast	N/A	N/A
WRN*	T573A	rs150148567	0.001	1	Colorectal, Breast, Pancreatic, Werner syndrome	DEL	Probably Damaging

Table 12. Rare missense variants in non-HBOC panel genes involved in DNA repair or cell cycle control and are associated with cancer phenotypes in HGMD. MAF=Minor Allele Frequency in (ExAC Non-Finnish Europeans.). DEL= Deleterious TOL= Tolerated, N/A= Information not available, *Variants in these genes were not confirmed by Sanger DNA sequencing.

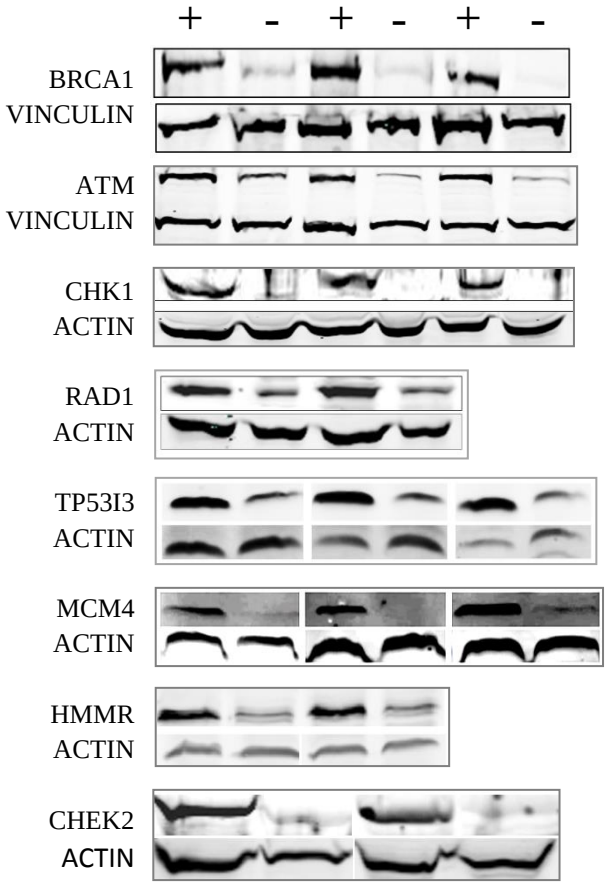


Figure 15. Western blots for siRNA knock down. Lysates from functional assays indicating protein expression with scramble (+) and reduction of same protein with targeted siRNA knock down (-) compared to Actin or Vinculin loading control.

REFERENCES

Akbari, M. R., Malekzadeh, R., Nasrollahzadeh, D., Amanian, D., Islami, F., Li, S., … Narod, S. a. (2008). Germline BRCA2 mutations and the risk of esophageal squamous cell carcinoma. *Oncogene*, *27*(9), 1290–1296. https://doi.org/10.1038/sj.onc.1210739

Al-Kaabi, M. M., Alshareeda, A. T., Jerjees, D. A., Muftah, A. A., Green, A. R., Alsubhi, N. H., … Rakha, E. A. (2015). Checkpoint kinase1 (CHK1) is an important biomarker in breast cancer having a role in chemotherapy response. *British Journal of Cancer*, *112*(5), 901–911. https://doi.org/10.1038/bjc.2014.576

American Cancer Society. (2014). *Cancer Facts & Figures*. *Cancer Facts and Figures*. Atlanta. Retrieved from http://www.cancer.org/acs/groups/content/@research/documents/webcontent/acspc-042151.pdf

Antoniou, A., Pharoah, P. D. P., Narod, S., Risch, H. A., Eyfjord, J. E., Hopper, J. L., … Easton, D. F. (2003). Average Risks of Breast and Ovarian Cancer-associated with BRCA1 or BRCA2 Mutations Detected in Case Series Unselected for Family History: A Combined Analysis of 22 Studies. *The American Journal of Human Genetics*, *72*(5), 1117–1130. https://doi.org/10.1086/375033

Bakker, S. T., van de Vrugt, H. J., Rooimans, M. A., Oostra, A. B., Steltenpool, J., Delzenne-Goette, E., … de Winter, J. P. (2009). Fancm-deficient mice reveal unique features of Fanconi anemia complementation group M. *Human Molecular Genetics*, *18*(18), 3484–3495. https://doi.org/10.1093/hmg/ddp297

Bansidhar, B. J. (2012). Extracolonic manifestations of lynch syndrome. *Clinics in Colon and Rectal Surgery*, *25*(2), 103–110. https://doi.org/10.1055/s-0032-1313781

Bast, R. C., Hennessy, B., & Mills, G. B. (2009). The biology of ovarian cancer: new opportunities for translation. *Nature Reviews. Cancer*, *9*(6), 415–428. https://doi.org/10.1038/nrc2644

Bell D, Berchuck A, Birrer M, Chien J, Cramer D, Dao F, Dhir R, DiSaia P, Gabra H, Glenn P, Godwin A, Gross J, Hartmann L, Huang M, Huntsman D, Iacocca M, Imielinski M, Kalloger S, Karlan B, Levine D, Mills G, Morrison C, Mutch D, Olvera N, Orsulic S, Par, T. E., & In, E. (2011). Integrated genomic analyses of ovarian carcinoma. *Nature*, *474*(7353), 609–615. https://doi.org/10.1038/nature10166

Bhatia, V., Barroso, S. I., García-Rubio, M. L., Tumini, E., Herrera-Moyano, E., & Aguilera, A. (2014). BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature*, *511*(7509), 362–365. https://doi.org/10.1038/nature13374

Bochar, D. a, Wang, L., Beniya, H., Kinev, a, Xue, Y., Lane, W. S., … Shiekhattar, R. (2000). BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer. *Cell*, *102*(2), 257–265. https://doi.org/10.1016/S0092-8674(00)00030-1

Bogliolo, M., & Surrallés, J. (2015). Fanconi anemia: a model disease for studies on human genetics and advanced therapeutics. *Current Opinion in Genetics & Development*, *33*, 32–40. https://doi.org/10.1016/j.gde.2015.07.002

Broeks, A., Urbanus, J. H., Floore, A. N., Dahler, E. C., Klijn, J. G., Rutgers, E. J., … van 't Veer, L. J. (2000). ATM-heterozygous germline mutations contribute to breast cancer-susceptibility. *American Journal of Human Genetics*, *66*(2), 494–500. https://doi.org/10.1086/302746

Bryant, H. E., Schultz, N., Thomas, H. D., Parker, K. M., Flower, D., Lopez, E., … Helleday, T.

(2005). Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature*, *434*(7035), 913–917. https://doi.org/10.1038/nature03443

Buisson, R., Dion-Côté, A.-M., Coulombe, Y., Launay, H., Cai, H., Stasiak, A. Z., … Masson, J.-Y. (2010). Cooperation of breast cancer proteins PALB2 and piccolo BRCA2 in stimulating homologous recombination. *Nat Struct Mol Biol.*, *17*(10), 1247–1254. https://doi.org/10.1002/nbm.3066.Non-invasive

Buisson, R., Niraj, J., Pauty, J., Maity, R., Zhao, W., Coulombe, Y., … Masson, J. Y. (2014). Breast cancer proteins PALB2 and BRCA2 stimulate polymerase η in recombination-associated DNA Synthesis At Blocked Replication Forks. *Cell Reports*, *6*(3), 553–564. https://doi.org/10.1016/j.celrep.2014.01.009

Centers for Disease Control and Prevention. (n.d.). Does Breast or Ovarian Cancer Run in Your Family? | Features | CDC. Retrieved December 13, 2017, from https://www.cdc.gov/features/hereditarycancer/

Chen, J. J., Silver, D., Cantor, S., Livingston, D. M., & Scully, R. (1999). BRCA1, BRCA2, and Rad51 operate in a common DNA damage response pathway. *Cancer Research*, *59*(7 SUPPL.), 1752–1757.

Chen, L., Nievera, C. J., Lee, A. Y.-L., & Wu, X. (2008). Cell Cycle-dependent Complex Formation of BRCA1·CtIP·MRN Is Important for DNA Double-strand Break Repair. *Journal of Biological Chemistry*, *283*(12), 7713–7720. https://doi.org/10.1074/jbc.M710245200

Chun, N., & Ford, J. M. (2012). Genetic Testing by Cancer Site. *The Cancer Journal*, *18*(4), 355–363. https://doi.org/10.1097/PPO.0b013e31826246dc

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., … Ruden, D. M. (2012).

A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w 1118; iso-2; iso-3. *Fly*, *6*(2), 80–92. https://doi.org/10.4161/fly.19695

Claus, E. B., Schildkraut, J. M., Thompson, W. D., & Risch, N. J. (1996). The genetic attributable risk of breast and ovarian cancer. *Cancer*, *77*(11), 2318–2324. https://doi.org/10.1002/(SICI)1097-0142(19960601)77:11<2318::AID-CNCR21>3.0.CO;2-Z

Clinical, N., Guidelines, P., & Guidelines, N. (2018). Genetic / Familial High-Risk Assessment : Breast and Ovarian.

DePristo, M. a., Banks, E., Poplin, R. E., Garimella, K. V., Maguire, J. R., Hartl, C., … Daly, M. J. (2011). A framework for variation discovery and genotyping using next- generation DNA sequencing data. *Nat Genet*, *43*(5), 491–498. https://doi.org/10.1038/ng.806.A

Domchek, S., & Weber, B. L. (2008a). Genetic variants of uncertain significance: Flies in the ointment. *Journal of Clinical Oncology*, *26*(1), 16–17. https://doi.org/10.1200/JCO.2007.14.4154

Domchek, S., & Weber, B. L. (2008b). Genetic variants of uncertain significance: flies in the ointment. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, *26*(1), 16–7. https://doi.org/10.1200/JCO.2007.14.4154

Erkko, H., Xia, B., Nikkilä, J., Schleutker, J., Syrjäkoski, K., Mannermaa, A., … Winqvist, R. (2007). A recurrent mutation in PALB2 in Finnish cancer families. *Nature*, *446*(7133), 316–319. https://doi.org/10.1038/nature05609

Evans, T., & Matulonis, U. (2017). PARP inhibitors in ovarian cancer: Evidence, experience and clinical potential. *Therapeutic Advances in Medical Oncology*, *9*(4), 253–267.

https://doi.org/10.1177/1758834016687254

Farmer, H., McCabe, N., Lord, C. J., Tutt, A. N. J., Johnson, D. A., Richardson, T. B., … Ashworth, A. (2005). Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature*, *434*(7035), 917–921. https://doi.org/10.1038/nature03445

Feliubadaló, L., Tonda, R., Gausachs, M., Trotta, J., Gel, B., Gut, M., … Menéndez, M. (2017). Benchmarking of Whole Exome Sequencing and Ad Hoc Designed Panels for Genetic Testing of Hereditary Cancer. *Nature Publishing Group*, (November 2016), 1–11. https://doi.org/10.1038/srep37984

Fletcher, O., Johnson, N., Silva, S., Orr, N., Ashworth, A., Nevanlinna, H., … Dunning, A. M. (2010). Europe PMC Funders Group Missense Variants in ATM in 26 , 101 Breast Cancer Cases and 29 , 842 Controls. *Cancer Epidemiol Biomarkers Prev*, *19*(9), 2143–2151. https://doi.org/10.1158/1055-9965.EPI-10-0374.Missense

Fleury, H., Communal, L., Carmona, E., Portelance, L., Arcand, S. L., Rahimi, K., … Mes-Masson, A.-M. (2015). Novel high-grade serous epithelial ovarian cancer cell lines that reflect the molecular diversity of both the sporadic and hereditary disease. *Genes & Cancer*, *6*(9–10), 378–398. https://doi.org/10.18632/genesandcancer.76

Frank, T. S. (1999). Testing for hereditary risk of ovarian cancer. *Cancer Control*, *6*(4), 327–324.

Gnad, F., Baucom, A., Mukhyala, K., Manning, G., & Zhang, Z. (2013). Assessment of computational methods for predicting the effects of missense mutations in human cancers. *BMC Genomics*, *14 Suppl 3*(Suppl 3), S7. https://doi.org/10.1186/1471-2164-14-S3-S7

Han, J., Hankinson, S. E., Vivo, I. De, Colditz, G. a, & Hunter, D. J. (2002). No Association between a Stop Codon Polymorphism in RAD52 and Breast Cancer Risk Null Results in

Brief No Association between a Stop Codon Polymorphism in RAD52 and. *Cancer Epidemiol Biomarkers Prev*, *11*(October), 1138–1139.

Han, L., Hu, Z., Liu, Y., Wang, X., Hopkins, K. M., Lieberman, H. B., & Hang, H. (2010). Mouse Rad1 deletion enhances susceptibility for skin tumor development. *Molecular Cancer*, *9*, 67. https://doi.org/10.1186/1476-4598-9-67

Helt, C. E., Cliby, W. A., Keng, P. C., Bambara, R. A., & O'Reilly, M. A. (2005). Ataxia telangiectasia mutated (ATM) and ATM and Rad3-related protein exhibit selective target specificities in response to different forms of DNA damage. *Journal of Biological Chemistry*, *280*(2), 1186–1192. https://doi.org/10.1074/jbc.M410873200

Hereditary Breast and Ovarian Cancer | Cancer.Net. (n.d.). Retrieved December 13, 2017, from https://www.cancer.net/cancer-types/hereditary-breast-and-ovarian-cancer

Janatova, M., Soukupova, J., Stribrna, J., Kleiblova, P., Vocka, M., Boudova, P., … Pohlreich, P. (2015). Mutation analysis of the RAD51C and RAD51D genes in high-risk ovarian cancer patients and families from the Czech Republic. *PLoS ONE*, *10*(6), 1–9. https://doi.org/10.1371/journal.pone.0127711

Jensen, R. B., Ozes, A., Kim, T., Estep, A., & Kowalczykowski, S. C. (2013). BRCA2 is epistatic to the RAD51 paralogs in response to DNA damage. *DNA Repair*, *12*(4), 306–311. https://doi.org/10.1016/j.dnarep.2012.12.007

Jin, S., & Weaver, D. T. (1997). Double-strand break repair by Ku70 requires heterodimerization with Ku80 and DNA binding functions. *The EMBO Journal*, *16*(22), 6874–6885. https://doi.org/10.1093/emboj/16.22.6874

Johnson, N., Fletcher, O., Palles, C., Rudd, M., Webb, E., Sellick, G., … Peto, J. (2007). Counting potentially functional variants in BRCA1, BRCA2 and ATM predicts breast

cancer susceptibility. *Human Molecular Genetics*, *16*(9), 1051–7. https://doi.org/10.1093/hmg/ddm050

Kaufman, B., Shapira-Frommer, R., Schmutzler, R. K., Audeh, M. W., Friedlander, M., Balmaña, J., … Domchek, S. M. (2015). Olaparib monotherapy in patients with advanced cancer and a germline BRCA1/2 mutation. *Journal of Clinical Oncology*, *33*(3), 244–250. https://doi.org/10.1200/JCO.2014.56.2728

Kim, G., Ison, G., McKee, A. E., Zhang, H., Tang, S., Gwise, T., … Pazdur, R. (2015). FDA approval summary: Olaparib monotherapy in patients with deleterious germline BRCA-mutated advanced ovarian cancer treated with three or more lines of chemotherapy. *Clinical Cancer Research*, *21*(19), 4257–4261. https://doi.org/10.1158/1078-0432.CCR-15-0887

Kim, H.-Y. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics*, *38*(1), 52. https://doi.org/10.5395/rde.2013.38.1.52

Kim, T. M., Son, M. Y., Dodds, S., Hu, L., & Hasty, P. (2014). Deletion of BRCA2 exon 27 causes defects in response to both stalled and collapsed replication forks. *Mutation Research*, *766–767*, 66–72. https://doi.org/10.1016/j.mrfmmm.2014.06.003.Deletion

Knudson, A. G. (1971). Mutation and Cancer: Statistical Study of Retinoblastoma. *Proceedings of the National Academy of Sciences*, *68*(4), 820–823. https://doi.org/10.1073/pnas.68.4.820

Kobayashi, H., Ohno, S., Sasaki, Y., & Matsuura, M. (2013). Hereditary breast and ovarian cancer susceptibility genes (Review). *Oncology Reports*, *30*(January), 1019–1029. https://doi.org/10.3892/or.2013.2541

Korch, C., Spillman, M. A., Jackson, T. A., Jacobsen, B. M., Murphy, S. K., Lesseye, B. A., … Bradford, A. P. (2012). DNA profiling analysis of endometrial and ovarian cell lines reveals

misidentification, redundancy and contamination. *Gynecologic Oncology*, *127*(1), 241–248. https://doi.org/10.1016/j.jacc.2007.01.076.White

Kumar, P., Al-Shafai, M., Al Muftah, W. A., Chalhoub, N., Elsaid, M. F., Aleem, A. A., & Suhre, K. (2014). Evaluation of SNP calling using single and multiple-sample calling algorithms by validation against array base genotyping and Mendelian inheritance. *BMC Research Notes*, *7*(1), 747. https://doi.org/10.1186/1756-0500-7-747

Laduca, H., Farwell, K. D., Vuong, H., Lu, H., Mu, W., Shahmirzadi, L., … Chao, E. C. (2017). Exome sequencing covers > 98 % of mutations identified on targeted next generation sequencing panels. *Plos One*, *12*(2), 1–11. https://doi.org/10.1371/journal.pone.0170843

Lamarche, B. J., Orazio, N. I., & Weitzman, M. D. (2010). The MRN complex in double-strand break repair and telomere maintenance. *FEBS Letters*, *584*(17), 3682–3695. https://doi.org/10.1016/j.febslet.2010.07.029

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., … International Human Genome Sequencing, C. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921. https://doi.org/10.1038/35057062

Lawrenson, K., Iversen, E. S., Tyrer, J., Weber, R. P., Concannon, P., Hazelett, D. J., … Schildkraut, J. M. (2015). Common variants at the CHEK2 gene locus and risk of epithelial ovarian cancer. *Carcinogenesis*, *36*(11), 1341–53. https://doi.org/10.1093/carcin/bgv138

Ledermann, J. A. (2016). PARP inhibitors in ovarian cancer. *Annals of Oncology*, *27*(December), i40–i44. https://doi.org/10.1093/annonc/mdw094

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liu, D., Shen, X., Zhu, G., & Xing, M. (2015). REC8 is a novel tumor suppressor gene epigenetically robustly targeted by the PI3K pathway in thyroid cancer. *Oncotarget*, *6*(36), 39211–24. https://doi.org/10.18632/oncotarget.5391

Liu, Q., Guntuku, S., Cui, X. S., Matsuoka, S., Cortez, D., Tamai, K., … Elledge, S. J. (2000). Chk1 is an essential kinase that is regulated by Atr and required for the G2/M DNA damage checkpoint. *Genes and Development*, *14*(12), 1448–1459. https://doi.org/10.1101/gad.14.12.1448

Lord, C. J., & Ashworth, A. (2017). PARP inhibitors: Synthetic lethality in the clinic. *Science*, *355*(March), 1152–1158.

Loveday, C., Turnbull, C., Ramsay, E., Hughes, D., Ruark, E., Frankum, J. R., … Davidson, R. (2016). Europe PMC Funders Group Germline mutations in RAD51D confer susceptibility to ovarian cancer. *Nature Genetics*, *43*(9), 879–882. https://doi.org/10.1038/ng.893.Germline

Ma, J.-L., Kim, E. M., Haber, J. E., & Lee, S. E. (2003). Yeast Mre11 and Rad1 Proteins Define a Ku-Independent Mechanism To Repair Double-Strand Breaks Lacking Overlapping End Sequences. *Molecular and Cellular Biology*, *23*(23), 8820–8828. https://doi.org/10.1128/MCB.23.23.8820-8828.2003

Mariotto, A. B., Robin Yabroff, K., Shao, Y., Feuer, E. J., & Brown, M. L. (2011). Projections of the cost of cancer care in the United States: 2010-2020. *Journal of the National Cancer Institute*, *103*(2), 117–128. https://doi.org/10.1093/jnci/djq495

Martin, S. T., Matsubayashi, H., Rogers, C. D., Philips, J., Couch, F. J., Brune, K., … Goggins, M. (2005). Increased prevalence of the BRCA2 polymorphic stop codon K3326X among individuals with familial pancreatic cancer. *Oncogene*, *24*(22), 3652–3656. https://doi.org/10.1038/sj.onc.1208411

Maxwell, C. A., Benítez, J., Gómez-Baldó, L., Osorio, A., Bonifaci, N., Fernández-Ramires, R., … Pujana, M. A. (2011). Interplay between BRCA1 and RHAMM regulates epithelial apicobasal polarization and may influence risk of breast cancer. *PLoS Biology*, *9*(11), 1–18. https://doi.org/10.1371/journal.pbio.1001199

Mehta, A., & Haber, J. E. (2014). Sources of DNA Double-Strand Breaks and Models of Rec. *Cold Spring Harbor Perspectives in Biology*, *6*, 1–19. https://doi.org/10.1101/cshperspect.a016428

Meindl, A., Hellebrand, H., Wiek, C., Erven, V., Wappenschmidt, B., Niederacher, D., … Hanenberg, H. (2010). Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. *Nature Genetics*, *42*(November 2009), 410–414. https://doi.org/10.1038/ng.569

Moisan, A., & Gaudreau, L. (2006). The BRCA1 COOH-terminal region acts as an RNA Polymerase II carboxyl-terminal domain kinase inhibitor that modulates p21WAF1/CIP1 expression. *Journal of Biological Chemistry*, *281*(30), 21119–21130. https://doi.org/10.1074/jbc.M600712200

Morak, M., Heidenreich, B., Keller, G., Hampel, H., Laner, A., de la Chapelle, A., & Holinski-Feder, E. (2014). Biallelic MUTYH mutations can mimic Lynch syndrome. *European Journal of Human Genetics : EJHG*, *22*(11), 1334–1337. https://doi.org/10.1038/ejhg.2014.15

Morimatsu, M., Donoho, G., & Hasty, P. (1998). Cells Deleted for Brca2 COOH Terminus Exhibit Hypersensitivity to y-Radiation and Premature Senescence. *Cancer Re*, *58*, 3441–3448. https://doi.org/10.1017/CBO9781107415324.004

Mu, K., Li, L., Yang, Q., Zhang, T., Gao, P., Meng, B., … Zhou, G. (2011). Detection of CHK1 and CCND1 gene copy number changes in breast cancer with dual-colour fluorescence in-situ hybridization. *Histopathology*, *58*(4), 601–607. https://doi.org/10.1111/j.1365-2559.2011.03780.x

Mucci, L. A., Hjelmborg, J. B., Harris, J. R., Czene, K., Havelick, D. J., Scheike, T., … Kaprio, J. (2016). Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. *Jama*, *315*(1), 68. https://doi.org/10.1001/jama.2015.17703

Mullany, L. K., Wong, K. K., Marciano, D. C., Katsonis, P., King-Crane, E. R., Ren, Y. A., … Richards, J. A. S. (2015). Specific TP53 Mutants Overrepresented in Ovarian Cancer Impact CNV, TP53 Activity, Responses to Nutlin-3a, and Cell Survival. *Neoplasia*, *17*(10), 789–803. https://doi.org/10.1016/j.neo.2015.10.003

Narod, S., Moody, J., Rosen, B., Fan, I., Risch, a, Sun, P., & McLaughlin, J. (2013). Estimating survival rates after ovarian cancer among women tested for BRCA1 and BRCA2 mutations. *Clinical Genetics*, *83*(3), 232–237. https://doi.org/10.1111/j.1399-0004.2012.01906.x

National Cancer Institute. (2013). *National Cancer Institute SEER Cancer Statistics Review 1975-2010*.

National Comprehensive Cancer Network. (2016). Genetic / Familial High-Risk Assessment: Colorectal 1.2016. *NCCN Guidelines*, 81. Retrieved from https://www.nccn.org/professionals/physician_gls/pdf/genetics_colon.pdf

National Comprehensive Cancer Network. Panel Members Genetic/Familial High-Risk

Assessment: Breast and Ovarian (Version 1.2018). (n.d.).

Ng, P. C. (2003). SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Research*, *31*(13), 3812–3814. https://doi.org/10.1093/nar/gkg509

Ngo, G. H. P., & Lydall, D. (2015). The 9-1-1 checkpoint clamp coordinates resection at DNA double strand breaks. *Nucleic Acids Research*, *43*(10), 5017–5032. https://doi.org/10.1093/nar/gkv409

NIH. (n.d.). Ovarian Cancer. Retrieved December 13, 2017, from https://medlineplus.gov/ovariancancer.html

NIH Surveillance, Epidemiology and End Results Program: Cancer of the Ovary - Cancer Stat Facts. (n.d.). Retrieved December 13, 2017, from https://seer.cancer.gov/statfacts/html/ovary.html

Norquist, B. M., Harrell, M. I., Brady, M. F., Walsh, T., Lee, M. K., Gulsuner, S., … Birrer, M. J. (2015). Inherited Mutations in Women With Ovarian Carcinoma. *JAMA Oncology*, *6460*(April 2016), 1–9. https://doi.org/10.1001/jamaoncol.2015.5495

Parmigiani, G., Berry, D., & Aguilar, O. (1998). Determining carrier probabilities for breast cancer-susceptibility genes BRCA1 and BRCA2. *American Journal of Human Genetics*, *62*(1), 145–58. https://doi.org/10.1086/301670

Paul D. P. Pharoah, Ya-Yu Tsai, Susan J. Ramus, Catherine M. Phelan, Ellen L. Goode, Kate Lawrenson, Melissa Price, Brooke L. Fridley, Jonathan P. Tyrer, Howard Shen, Rachel Weber, Rod Karevan, Melissa C. Larson, Honglin Song, Daniel C. Tessier, François, T. A. S. (2013). GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. *Nature Genetics*, *45*(4), 362–370e2.

PDQ Cancer Genetics Editorial Board, P. C. G. E. (2002). *Genetics of Breast and Gynecologic*

Cancers (PDQ®): Health Professional Version. *PDQ Cancer Information Summaries*. National Cancer Institute (US). Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/26389210

Perego, P., Zunino, F., Carenini, N., Giuliani, F., Spinelli, S., & Howell, S. B. (1998). Sensitivity to cisplatin and platinum-containing compounds of Schizosaccharomyces pombe rad mutants. *Molecular Pharmacology*, *54*(1), 213–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9658208

Peterlongo, P., Catucci, I., Colombo, M., Caleca, L., Mucaki, E., Bogliolo, M., … Radice, P. (2015). *FANCM* c.5791C>T nonsense mutation (rs144567652) induces exon skipping, affects DNA repair activity and is a familial breast cancer risk factor. *Human Molecular Genetics*, ddv251. https://doi.org/10.1093/hmg/ddv251

Ph, D., & Andrea, A. D. D. (2016). *Fundamental Vulnerability of Ovarian Cancer* (Vol. 5). https://doi.org/10.1158/2159-8290.CD-15-0714.Homologous

Plevová, P., Novotný, J., Petráková, K., Palácová, M., Kalábová, R., Schneiderová, M., & Foretová, L. (2009). Hereditary breast and ovarian cancer syndrome. *Klinicka Onkologie : Casopis Ceske a Slovenske Onkologicke Spolecnosti*, *22 Suppl*, S8–S11.

Povirk, L. F. (2012). Processing of Damaged DNA Ends for Double-Strand Break Repair in Mammalian Cells. *ISRN Molecular Biology*, *2012*, 1–16. https://doi.org/10.5402/2012/345805

Pujana, M. A., Han, J.-D. J., Starita, L. M., Stevens, K. N., Tewari, M., Ahn, J. S., … Vidal, M. (2007). Network modeling links breast cancer susceptibility and centrosome dysfunction. *Nature Genetics*, *39*(11), 1338–1349. https://doi.org/10.1038/ng.2007.2

Rahman, N., Seal, S., Thompson, D., Kelly, P., Renwick, A., Elliott, A., … Stratton, M. R.

(2010). UKPMC Funders Group cancer susceptibility gene. *Cancer*, *39*(2), 165–167. https://doi.org/10.1038/ng1959.PALB2

Ramus, S. J., & Gayther, S. A. (2009). The Contribution of BRCA1 and BRCA2 to Ovarian Cancer. *Molecular Oncology*, *3*(2), 138–150. https://doi.org/10.1016/j.molonc.2009.02.001

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., … Rehm, H. L. (2015a). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, *17*(5), 405–424. https://doi.org/10.1038/gim.2015.30

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., … Rehm, H. L. (2015b). Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Medical Genetics*, *17*(5), 405–424. https://doi.org/10.1038/gim.2015.30.Standards

Rijkers, T., Van Den Ouweland, J., Morolli, B., Rolink, A. G., Baarends, W. M., Van Sloun, P. P. H., … Pastink, A. (1998). Targeted Inactivation of Mouse *RAD52* Reduces Homologous Recombination but Not Resistance to Ionizing Radiation. *Molecular and Cellular Biology*, *18*(11), 6423–6429. https://doi.org/10.1128/MCB.18.11.6423

Risch, H. A., McLaughlin, J. R., Cole, D. E. C., Rosen, B., Bradley, L., Kwan, E., … Narod, S. A. (2001). Prevalence and Penetrance of Germline BRCA1 and BRCA2 Mutations in a Population Series of 649 Women with Ovarian Cancer. *The American Journal of Human Genetics*, *68*(3), 700–710. https://doi.org/10.1086/318787

Risk Factors - Ovarian Cancer Research Fund Alliance. (n.d.). Retrieved December 13, 2017,

from https://ocrfa.org/patients/about-ovarian-cancer/risk-factors/

Robles-Díaz, L., Goldfrank, D., Kauff, N., Robson, M., & Offit, K. (2004). Hereditary ovarian cancer in Ashkenazi Jews. *Familial Cancer*, *3*(3–4), 259–264. https://doi.org/10.1007/s10689-004-9552-0

Robson, M., Im, S.-A., Senkus, E., Xu, B., Domchek, S. M., Masuda, N., … Conte, P. (2017). Olaparib for Metastatic Breast Cancer in Patients with a Germline *BRCA* Mutation. *New England Journal of Medicine*, *377*(6), 523–533. https://doi.org/10.1056/NEJMoa1706450

Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S., & Bonner, W. M. (1998). Double-stranded Brekas Induce Histone H2AX phosphorylation on Serine 139. *The Journal of Biological Chemistry*, *273*(10), 5858–5868. https://doi.org/10.1074/jbc.273.10.5858

Rudd, M. F., Webb, E. L., Matakidou, A., Sellick, G. S., Williams, R. D., Bridle, H., … Consortium, G. (2006). Variants in the GH-IGF axis confer susceptibility to lung cancer. *Genome Res*, *16*(6), 693–701. https://doi.org/10.1101/gr.5120106

Ryland, G. L., Doyle, M. A., Goode, D., Boyle, S. E., Choong, D. Y. H., Rowley, S. M., … Gorringe, K. L. (2015). Loss of heterozygosity: What is it good for? *BMC Medical Genomics*, *8*(1), 1–12. https://doi.org/10.1186/s12920-015-0123-z

Schwab, R. A., Nieminuszczy, J., Shah, F., Langton, J., Lopez Martinez, D., Liang, C. C., … Niedzwiedz, W. (2015). The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication and Transcription. *Molecular Cell*, *60*(3), 351–361. https://doi.org/10.1016/j.molcel.2015.09.012

Sekhar, C., Chilamakuri, R., Lorenz, S., Madoui, M., Vodák, D., Sun, J., … Meza-zepeda, L. A. (2014). Performance comparison of four exome capture systems for deep sequencing Performance comparison of four exome capture systems for deep sequencing.

Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., & Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research*, *29*(1), 308–11. https://doi.org/10.1093/nar/29.1.308

Sheu, Y.-J., Kinney, J. B., Lengronne, A., Pasero, P., & Stillman, B. (2014). Domain within the helicase subunit Mcm4 integrates multiple kinase signals to control DNA replication initiation and fork progression. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(18), E1899-908. https://doi.org/10.1073/pnas.1404063111

Shima, N., Alcaraz, A., Liachko, I., Buske, T. R., Andrews, C. A., Munroe, R. J., … Schimenti, J. C. (2007). A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nature Genetics*, *39*(1), 93–8. https://doi.org/10.1038/ng1936

Sneeden, J. L., Grossi, S. M., Tappin, I., Hurwitz, J., & Heyer, W. D. (2013). Reconstitution of recombination-associated DNA synthesis with human proteins. *Nucleic Acids Research*, *41*(9), 4913–4925. https://doi.org/10.1093/nar/gkt192

South, C. D., Hampel, H., Comeras, I., Westman, J. A., Frankel, W. L., & De La Chapelle, A. (2008). The frequency of Muir-Torre syndrome among Lynch syndrome families. *Journal of the National Cancer Institute*, *100*(4), 277–281. https://doi.org/10.1093/jnci/djm291

Spring, C., & Perspect, H. (2015). Homologous recombination and human health: the roles of BRCA1, BRCA2 and associated proteins. *Cold Spring Harb Perspec Biol*, *7*(4), 1–37. https://doi.org/10.1101/cshperspect.a016600.Homologous

Stafford, J. L., Dyson, G., Levin, N. K., Chaudhry, S., Rosati, R., Kalpage, H., … Tainsky, M. A. (2017). Reanalysis of BRCA1/2 negative high risk ovarian cancer patients reveals novel germline risk loci and insights into missing heritability. *Plos One*, *12*(6), e0178450. https://doi.org/10.1371/journal.pone.0178450

Stancel, N. J., McDaniel, L. D., Velasco, S., Richardson, J., Guo, C., & Friedberg, E. C. (2009). Polk Mutant Mice Have a Spontaneous Mutator Phenotype. *DNA Repair (Amst).*, *8*(12), 1355–1362. https://doi.org/10.1016/j.freeradbiomed.2008.10.025.The

Stanislaw C., Xue Y., Wilcox WR. Genetic evaluation and testing for hereditary forms of cancer in the era of next-generation sequencing. Cancer Biol Med. 2016; 13: 55-67. doi: 10.28092/j.issn.2095- 3941.2016.0002

Stenson, P. D., Ball, E. V., Mort, M., Phillips, A. D., Shiel, J. A., Thomas, N. S. T., … Cooper, D. N. (2003). Human Gene Mutation Database (HGMD): 2003 Update. *Human Mutation*, *21*(6), 577–581. https://doi.org/10.1002/humu.10212

Strafford, J. C. (2012). Genetic testing for lynch syndrome, an inherited cancer of the bowel, endometrium, and ovary. *Reviews in Obstetrics & Gynecology*, *5*(1), 42–9. https://doi.org/10.3909/riog0187

Sy, S. M. H., Huen, M. S. Y., Zhu, Y., & Chen, J. (2009). PALB2 regulates recombinational repair through chromatin association and oligomerization. *Journal of Biological Chemistry*, *284*(27), 18302–18310. https://doi.org/10.1074/jbc.M109.016717

Tarsounas, M., Davies, A. A., & West, S. C. (2004). RAD51 localization and activation following DNA damage. *Philisophical Transactions of the Royal Society B: Biological Sciences*, *349*(1441), 87–93. https://doi.org/10.1098/rstb.2003.1368

The CHEK2 Breast Cancer Case-Control Consortium. (2004). CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. *American Journal of Human Genetics*, *74*(6), 1175–1182. https://doi.org/10.1086/421251

Thompson, E. R., Gorringe, K. L., Rowley, S. M., Li, N., McInerny, S., Wong-Brown, M. W., …

Campbell, I. G. (2015). Reevaluation of the BRCA2 truncating allele c.9976A > T (p.Lys3326Ter) in a familial breast cancer context. *Scientific Reports*, *5*(April 2015), 14800. https://doi.org/10.1038/srep14800

Thompson, E. R., Rowley, S. M., Sawyer, S., Eccles, D. M., Trainer, A. H., Mitchell, G., … Campbell, I. G. (2013). Analysis of RAD51D in Ovarian Cancer Patients and Families with a History of Ovarian or Breast Cancer. *PLoS ONE*, *8*(1). https://doi.org/10.1371/journal.pone.0054772

Thorstenson, Y. R., Roxas, A., Kroiss, R., Jenkins, M. A., Yu, K. M., Bachrich, T., … Oefner, P. J. (2003). Contributions of ATM Mutations to Familial Breast and Ovarian Cancer 1, 3325–3333.

Towler, W. I., Zhang, J., Ransburgh, D. J. R., Toland, A. E., Ishioka, C., Chiba, N., & Parvin, J. D. (2013). Analysis of BRCA1 Variants in Double-Strand Break Repair by Homologous Recombination and Single-Strand Annealing. *Human Mutation*, *34*(3), 439–445. https://doi.org/10.1002/humu.22251

Treszezamsky, A. D., Kachnic, L. A., Feng, Z., Zhang, J., Tokadjian, C., & Powell, S. N. (2007). BRCA1- and BRCA2-deficient cells are sensitive to etoposide-induced DNA double-strand breaks via topoisomerase II. *Cancer Research*, *67*(15), 7078–7081. https://doi.org/10.1158/0008-5472.CAN-07-0601

Trifonov, I., Todorova, M., & Uzunova, Z. (2001). Hereditary ovarian cancer. *Akusherstvo I Ginekologiia*, *41 Suppl 4*, 3–7. https://doi.org/10.1016/0165-4608(91)90094-B

van Gent, D. C., & Kanaar, R. (2016). Exploiting DNA repair defects for novel cancer therapies. *Molecular Biology of the Cell*, *27*(14), 2145–2148. https://doi.org/10.1091/mbc.E15-10-0698

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *American Journal of Human Genetics*, *90*(1), 7–24. https://doi.org/10.1016/j.ajhg.2011.11.029

Walsh, T., Casadei, S., Lee, M. K., Pennil, C. C., Nord, a. S., Thornton, a. M., … Swisher, E. M. (2011). From the Cover: Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. *Proceedings of the National Academy of Sciences*, *108*(44), 18032–18037. https://doi.org/10.1073/pnas.1115052108

Wang, H. F., Takenaka, K., Nakanishi, A., & Miki, Y. (2011). BRCA2 and nucleophosmin coregulate centrosome amplification and form a complex with the Rho effector kinase ROCK2. *Cancer Research*, *71*(1), 68–77. https://doi.org/10.1158/0008-5472.CAN-10-0030

Wang, Y., Huang, B., Edelmann, L., Kolodner, R. D., & Edelmann, W. (2010). Genes essential for cell viability that are linked to tumor suppressor genes play a role in cancer susceptibility. *Cell Cycle (Georgetown, Tex.)*, *9*(16), 3144–5. https://doi.org/10.4161/cc.9.16.12951

Wang, Y., McKay, J. D., Rafnar, T., Wang, Z., Timofeeva, M., Broderick, P., … Valk, K. (2015). Europe PMC Funders Group Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer, *46*(7), 736–741. https://doi.org/10.1038/ng.3002.Rare

Wu, K., Hinson, S. S. R., Ohashi, A., Farrugia, D., Wendt, P., Tavtigian, S. V, … Couch, F. J. (2005). Functional evaluation and cancer risk assessment of BRCA2 unclassified variants. *Cancer Research*, *65*(February 2005), 417–426. Retrieved from http://cancerres.aacrjournals.org/content/65/2/417.short

Wu, W., Koike, A., Takeshita, T., & Ohta, T. (2008). The ubiquitin E3 ligase activity of BRCA1

and its biological functions. *Cell Division*, *3*(1), 1. https://doi.org/10.1186/1747-1028-3-1

Xu, H., Beasley, M. D., Warren, W. D., van der Horst, G. T. J., & McKay, M. J. (2005). Absence of mouse REC8 cohesin promotes synapsis of sister chromatids in meiosis. *Developmental Cell*, *8*(6), 949–961. https://doi.org/10.1016/j.devcel.2005.03.018

Yang, J., Bogni, A., Schuetz, E. G., Ratain, M., Eileen, M., Dolan, H., … Altmand. (2009). Etoposide pathway. *Pharmacogenet Genomics*, *19*(7). https://doi.org/10.3174/ajnr.A1256.Functional

Yarden, R. I., Metsuyanim, S., Pickholtz, I., Shabbeer, S., Tellio, H., & Papa, M. Z. (2012). BRCA1-dependent Chk1 phosphorylation triggers partial chromatin disassociation of phosphorylated Chk1 and facilitates S-phase cell cycle arrest. *International Journal of Biochemistry and Cell Biology*, *44*(11), 1761–1769. https://doi.org/10.1016/j.biocel.2012.06.026

Yoon, S. W., Lee, M. S., Xaver, M., Zhang, L., Hong, S. G., Kong, Y. J., … Kim, K. P. (2016). Meiotic prophase roles of Rec8 in crossover recombination and chromosome structure. *Nucleic Acids Research*, *44*(19), 9296–9314. https://doi.org/10.1093/nar/gkw682

Yoshida, K., & Miki, Y. (2004). Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. *Cancer Science*, *95*(11), 866–871. https://doi.org/10.1111/j.1349-7006.2004.tb02195.x

Yousefzadeh, M. J., & Wood, R. D. (2013). DNA polymerase POLQ and cellular defense against DNA damage. *DNA Repair*, *12*(1), 1–9. https://doi.org/10.1016/j.dnarep.2012.10.004

Yu, J., Liang, Q., Wang, J., Wang, K., Gao, J., Zhang, J., … Sung, J. J. Y. (2017). REC8 functions as a tumor suppressor and is epigenetically downregulated in gastric cancer, especially in EBV-positive subtype. *Oncogene*, *36*(2), 182–193.

https://doi.org/10.1038/onc.2016.187

Zhang, W., Luo, J., Chen, F., Yang, F., & Song, W. (2015). BRCA1 regulates PIG3-mediated apoptosis in a p53-dependent manner. *Oncotarget*, *6*(10), 7606–7618.

ABSTRACT

MISSING HERITABILITY AND NOVEL GERMLINE RISK LOCI IN HEREDITARY OVARIAN CANCER: INSIGHTS FROM WHOLE EXOME SEQUENCING AND FUNCTIONAL ANALYSES

by

JAIME L. STAFFORD

May 2018

Advisor: Dr. Michael A. Tainsky

Major: Molecular Genetics and Genomics

Degree: Doctor of Philosophy

While 25% of ovarian cancer (OVCA) cases are due to inherited factors, most of the genetic risk remains unexplained. This study addressed this gap by identifying previously undescribed OVCA risk loci through the whole exome sequencing (WES) of 48 *BRCA1/BRCA2* wild type women diagnosed with OVCA, selected for high risk of genetic inheritance. Five clearly pathogenic variants were identified in this sample, four of which are in two genes featured on current multi-gene panels; (*RAD51D, ATM*). In addition, a high impact variant in *FANCM* (R1931*) was identified. *FANCM* has been recently implicated in familial breast cancer risk but is not yet featured on testing panels. Numerous rare and predicted to be damaging variants of unknown significance were detected in genes on current commercial testing panels. Also, the *BRCA2* variant p.K3326*, considered benign but resulting in a 93 amino acid truncation, was overrepresented in our sample (OR= 4.95, p=0.01) and coexisted in the germline of these women with other deleterious variants, suggesting a possible role as a modifier of genetic penetrance.

A candidate gene analysis detected loss of function (LOF) variants in genes involved in OVCA relevant pathways; DNA repair and cell cycle control, including *FANCM*, *CHK1, TP53I3, REC8, HMMR, RAD1,* and *MCM4*. Wet lab functional assessment implicated *FANCM, CHK1, RAD1* and *TP53I3* as having the *BRCA*-like phenotype typically observed in tumor suppressor genes commonly mutated the germline of women with inherited risk of breast and/or ovarian cancer. Importantly, plotting various panel genes based on cell viability and sensitivity to DNA damage after siRNA knock down correctly differentiated between moderate and high penetrant genes. This technique identified candidate genes *CHK1* and *RAD1* as high and *TP53I3* as moderate in penetrance.

The results of this project indicate that WES on study samples filtered for family history and negative for known causal variants is the most appropriate study design for identifying rare and novel high-risk variants. This study implicates novel risk loci as well as highlights the necessity of wet lab functional assessment. Importantly, this study also suggests that wet lab assays may be employed to differentiate moderate from high risk genetic loci.

AUTOBIOGRAPHICAL STATEMENT

JAIME L. STAFFORD

Education

2014-2018	Ph.D.	Wayne State University, School of Medicine, Detroit MI Major: Molecular Genetics and Genomics
2012-2013	M.S	University College London, London, United Kingdom, Major: Genetics of Human Disease
2002-2005	B.A	Oakland University, Rochester, MI, Major: Anthropology

Publications

1. **Stafford J.L**., (2018) Genetic Predispositions. In: Shackelford T., Weekes-Shackelford V. (eds) Encyclopedia of Evolutionary Psychological Science. Springer, Cham
2. **Stafford J.L**., Chaudhry S.R, Levin N.K, Tainsky M.A. (2017) Whole exome sequencing: a necessary tool for the future of clinical cancer care. *J Cancer Biol Res 5(3): 1106.*
3. **Stafford, J.L**., Dyson G., Levin N.K., Chaudhry S., Rosati R., Kalpage H., Wernette C., Petrucelli N., Simon M.S., and Tainsky M.A. (2017). "Reanalysis of BRCA1/2 Negative High Risk Ovarian Cancer Patients Reveals Novel Germline Risk Loci and Insights into Missing Heritability." PlosOne 12 (6): e0178450. doi:10.1371/journal.pone.0178450. PMID: 28591191
4. Demetriou, C., S. Abu-Amero, A.C. Thomas, M. Ishida, R. Aggarwal, L. Al-Olabi, L.J. Leon, **Stafford, J L**., et al. 2014. "Paternally Expressed, Imprinted Insulin-like Growth Factor-2 in Chorionic Villi Correlates Significantly with Birth Weight." PLoS ONE9 (1). doi:10.1371/journal.pone.0085454. PMID: 24454871
5. Moore, G.E., M. Ishida, C. Demetriou, L. Al-Olabi, L.J. Leon, A.C. Thomas, S. Abu-Amero, **Stafford, J L**., et al. 2015. "The Role and Interaction of Imprinted Genes in Human Fetal Growth." Philosophical Transactions of the Royale Society B: Biological Sciences370 (1663). doi:10.1098/rstb.2014.0074. PMID: 25602077

Awards and Honors

Thomas C. Rumble Fellowship University Graduate Fellowship 2017-2018
Graduate and Post-Doctoral Research Symposium, Wayne State, Poster, Third Place,
2017 Graduate Student Research Day, Wayne State, First Place Poster Presentation, 2016
ACMG Top Rated Abstract, 2016
NIH Broadening Experiences in Scientific Training Award (BEST) 2016
Interdisciplinary Biomedical Science Fellowship, 2014
Award for "Best Thesis" UCL, 2013
Highest Academic Honors Award GPA 3.9, 2013
UCL Master Prize Scholarship, 2012