

ANALYSIS OF MATCHED PRIMARY AND RECURRENT *BRCA1/2* MUTATION-
ASSOCIATED TUMORS IDENTIFIES RECURRENCE-SPECIFIC DRIVERS

Jennifer Brady Shah

A DISSERTATION

in

Cell and Molecular Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2021

Supervisor of Dissertation

Katherine L. Nathanson, MD, Professor of Medicine

Graduate Group Chairperson

Daniel S. Kessler, PhD, Associate Professor of Cell and Developmental Biology

Dissertation Committee

Roger A. Greenberg, MD, PhD, Professor of Cancer Biology

Ronny Drapkin, MD, PhD, Associate Professor of Obstetrics and Gynecology

Marylyn D. Ritchie, PhD, Professor of Genetics

Kai Tan, PhD, Professor of Pediatrics

ANALYSIS OF MATCHED PRIMARY AND RECURRENT *BRCA1/2* MUTATION-
ASSOCIATED TUMORS IDENTIFIES RECURRENCE-SPECIFIC DRIVERS

COPYRIGHT

2021

Jennifer Brady Shah

This work is licensed under the
Creative Commons Attribution-
NonCommercial-ShareAlike 4.0 International
License

To view a copy of this license, visit

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

To my parents, who got me here

ACKNOWLEDGMENTS

First, I would like to thank the patients included in this study, and their families, for the use of biological specimens and clinical data. Unfortunately, most of the patients in this cohort are now deceased, but we hope that their generosity will improve outcomes for *BRCA1/2* mutation carriers in the future. I would also like to thank Nathanson Lab members and lab alumni for their help with this project, including John Pluta, Kurt D'Andrea, Brad Wubbenhorst, Liza Dorfman, Anna Hubert, Dr. Adam Kraya, Dr. Dana Pueschl, and Dr. Kara Maxwell. I would like to thank my advisor Dr. Kate Nathanson, for her continued mentorship and guidance, and for the freedom to take this project in unexpected directions. I benefitted immensely from the expertise of my thesis committee (Drs. Marylyn Ritchie, Ronny Drapkin, Roger Greenberg, and Kai Tan), who provided thoughtful feedback throughout the project. Lastly, I would like to thank my program leadership (and especially Drs. Sandra Ryeom, Craig Bassing, and Dan Kessler) for their mentorship and advocacy during my tenure as a graduate student.

Next, I would like to thank the many friends who supported me throughout my time in grad school. In particular, I should thank Gloria, Sara, Rebekah, and Aurora for helping me to stay positive (and commiserating with me when I could not). To all of the labmates-turned-friends I made along the way (especially Gabby, Trish, and Jake), thanks to you as well. I would also like to thank my numerous non-science friends from high school, college, and our time in Philly. I am incredibly lucky to have met such great people, who continue to remind me that life goes on outside of work. The last six years would have gone by much more slowly without you all!

I could not have completed this degree without the support of my family. I would like to thank my parents and brother for their endless love and support (and for positing that their PhD's were probably more difficult than mine). I would like to thank Amma and Pappa, who (thankfully) asked me about what I was cooking and how far I was running more than they asked about work. I would like to thank my Auntie Parul and Uncle Ashish, who let me vent when I needed it most. Lastly, I must acknowledge my cat Muffin, who reminds me to enjoy the little things (like

birdwatching and naps in the sunshine). Completing a PhD during a pandemic wasn't easy, but you all kept me sane and grounded this year.

Last but not least, I would like to thank my partner Andres, who has supported me throughout my graduate school career. Andres left a job and a city he loved to join me in Philly, and he has seen me through every curveball since. For someone with a high school biology background, he now knows more about cancer genetics than most. Thank you for listening to me, making me laugh, and taking me on all those walks. Most of all, thanks for reminding me that I am a person first and a scientist second (or maybe even third).

ABSTRACT

ANALYSIS OF MATCHED PRIMARY AND RECURRENT *BRCA1/2* MUTATION-ASSOCIATED TUMORS IDENTIFIES RECURRENCE-SPECIFIC DRIVERS

Jennifer B. Shah

Katherine L. Nathanson

Patients with inherited germline mutations in *BRCA1/2* carry a drastically increased risk of early-onset breast and ovarian cancers. *BRCA1/2* mutation-associated tumors respond to therapies that exploit their inherent homologous recombination deficiency, including platinum-based chemotherapy and poly(ADP-ribose) polymerase inhibitors (PARPi). However, these tumors frequently return as lethal, therapy-resistant recurrences. Outside of somatic *BRCA1/2* reversions, the mechanisms underlying acquired therapeutic resistance and recurrence remain unknown. To address this gap in knowledge, we performed whole exome, targeted, and RNA sequencing on paired primary and recurrent breast and ovarian tumors from 27 *BRCA1/2* mutation carriers. The main outcomes of the study were somatic variants, copy number variation, *BRCA1/2* loss of heterozygosity, differential gene expression, and differential transcript usage. One key finding was a high prevalence of copy number gains and amplifications in *PARP1*. We detected *PARP1* gains across primary and recurrent *BRCA1/2* mutation-associated breast and ovarian tumors, with increased expression observed at the mRNA and protein levels. Our results suggest that *PARP1* gains could be an under-appreciated mechanism of endogenous PARPi resistance in *BRCA1/2* mutation-associated and sporadic breast and ovarian tumors. We also assessed allele-specific *BRCA1/2* loss of heterozygosity (LOH) across the cohort. In general, LOH status was concordant between paired primary and recurrent tumors. However, seven tumors underwent LOH

transitions over the course of recurrence, suggesting that selective pressure drove most cancers towards biallelic *BRCA1/2* loss but some towards *BRCA1/2* proficiency. Lastly, we found that recurrent tumors express a shorter *BRCA2* transcript. This non-canonical isoform is protein-coding and differs only in the 3' UTR. Expression of the alternative *BRCA2* transcript was significantly associated with reduced overall survival in *BRCA1/2* mutation carriers with breast cancer (median 87 vs. 121 months). Our results indicate that the shorter *BRCA2* isoform may represent a novel driver of recurrence in *BRCA1/2* mutation-associated breast tumors. Ultimately, these findings improve our understanding of tumor evolution in *BRCA1/2* mutation-associated cancers, including conserved drivers and other features that may contribute to therapeutic resistance. Ultimately, this work will improve our understanding of late-stage disease in *BRCA1/2* mutation carriers, as well as inform new treatment options for these patients.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	IV
ABSTRACT	VI
LIST OF TABLES	XI
LIST OF ILLUSTRATIONS	XII
LIST OF ABBREVIATIONS	XV
CHAPTER 1: INTRODUCTION	1
<i>BRCA1/2 Proteins in DNA Repair</i>	1
<i>BRCA1/2 Mutations and Cancer Presentation</i>	2
<i>BRCA1/2 Loss of Heterozygosity</i>	4
<i>Acquired Therapeutic Resistance in BRCA1/2 Mutation-associated Tumors</i>	5
<i>Remaining Gaps in Knowledge</i>	8
CHAPTER 2: PATIENT COHORT AND MULTI-OMIC SEQUENCING	11
<i>Introduction</i>	11
<i>Methods</i>	11
<i>Results</i>	15
<i>Discussion</i>	23
CHAPTER 3: ANALYSIS OF SOMATIC MUTATIONS	25
<i>Introduction</i>	25
<i>Methods</i>	26
<i>Results</i>	29
<i>Discussion</i>	39
CHAPTER 4: ANALYSIS OF COPY NUMBER VARIATION	43

<i>Introduction</i>	43
<i>Methods</i>	45
<i>Results</i>	48
<i>Discussion</i>	67
CHAPTER 5: DIFFERENTIAL GENE EXPRESSION AND GENE FUSIONS ...	71
<i>Introduction</i>	71
<i>Methods</i>	72
<i>Results</i>	75
<i>Discussion</i>	93
CHAPTER 6: DIFFERENTIAL TRANSCRIPT USAGE	96
<i>Introduction</i>	96
<i>Methods</i>	97
<i>Results</i>	99
<i>Discussion</i>	108
CHAPTER 7: CASE STUDIES IN THREE PATIENTS	110
<i>Introduction</i>	110
<i>Methods</i>	110
<i>Results</i>	110
<i>Discussion</i>	117
CHAPTER 8: CONCLUSIONS AND FUTURE DIRECTIONS	120
<i>Conclusions</i>	120
<i>Limitations</i>	124
<i>Future Directions (Bench Experiments)</i>	125
<i>Future Directions (Bioinformatics)</i>	128
<i>Data and Code Availability</i>	130

APPENDIX	131
BIBLIOGRAPHY	132

LIST OF TABLES

Table 1: Patient Cohort and Tumor Characteristics

Table 2: Genes Included in Targeted Capture

LIST OF ILLUSTRATIONS

- Figure 1: Distribution of recurrences collected by anatomical site
- Figure 2: Multi-omic sequencing completed by tumor
- Figure 3: Overview of analyses from whole exome sequencing
- Figure 4: Overview of analyses from RNA sequencing
- Figure 5: Manual filtering strategy for somatic variants
- Figure 6: Key mutations detected by whole exome and targeted sequencing
- Figure 7: Somatic mutations and mutational signatures by tumor
- Figure 8: Paired comparisons of tumor mutational burden
- Figure 9: Pathway analysis of loss of function mutations
- Figure 10: Significant mutations in individual genes
- Figure 11: Filtering strategy for TCGA PanCancer cohorts
- Figure 12: *BRCA1/2* loss of heterozygosity by tumor
- Figure 13 : *BRCA1* germline variants, B allele frequency, and copy number before and after a representative nonLOH to LOH transition
- Figure 14: *BRCA1* germline variants, B allele frequency, and copy number before and after a representative LOH reversal
- Figure 15: HRD score comparisons in tumors with *BRCA1/2* LOH transitions
- Figure 16: Comparisons of HRD and aneuploidy scores in primary/recurrent tumors
- Figure 17: Copy number amplifications (GISTIC) in breast and ovarian tumors
- Figure 18: Copy number deletions (GISTIC) in breast and ovarian tumors
- Figure 19: Copy number amplifications (GISTIC) in primary and recurrent tumors
- Figure 20: Copy number deletions (GISTIC) in primary and recurrent tumors
- Figure 21: *PARP1* copy number variation in primary/recurrent cohort
- Figure 22: Segments of *PARP1* gains and amplifications by tumor
- Figure 23: *PARP1* copy number variation in TCGA PanCancer cohorts
- Figure 24: Pathway analysis of genes subject to copy number variation
- Figure 25: Distribution of $\text{Log}_2(\text{counts per million})$ across cohort of RNA-seq samples

Figure 26: Cohort of 66 samples used for RNA sequencing

Figure 27: Dendrogram from agglomerative hierarchical clustering of RNA sequencing cohort

Figure 28: Principal components analysis of RNA-seq cohort

Figure 29: Differential gene expression in recurrent tumors vs. normal tissue

Figure 30: Differential gene expression in primary tumors vs. normal tissue

Figure 31: Pathway analysis of over-expressed genes in tumors

Figure 32: Pathway analysis of over-expressed genes in tumors by tumor type

Figure 33: Identification and functional enrichment of gene modules based on tumor vs. normal expression differences

Figure 34: Functional enrichment of gene modules

Figure 35: Gene set enrichment analysis in primary vs. recurrent tumors

Figure 36: *PARP1* mRNA expression in primary and recurrent tumors

Figure 37: PARP1 protein expression in primary and recurrent tumors

Figure 38: Gene fusions detected by RNA sequencing

Figure 39: *BRCA2* gene and transcript expression in normal tissue and recurrent tumors

Figure 40: *BRCA2* isoform expression by sample and group for entire RNA sequencing cohort

Figure 41: RNA binding proteins in the 3' UTR region of difference between *BRCA2* transcripts

Figure 42: Validation of *BRCA2* isoform detection in an independent cohort

Figure 43: Effect of differential *BRCA2* transcript usage on overall survival in *BRCA1/2* mutation carriers with breast cancer

Figure 44: Effect of differential *BRCA2* transcript usage on overall survival in *BRCA1/2* mutation carriers with ovarian cancer

Figure 45: Genetic and clinical features in Patient 20, a *BRCA1* mutation carrier with ovarian cancer

Figure 46: Somatic reversion of *BRCA1* in Patient 20

Figure 47: Sequencing tracks demonstrating *BRCA1* reversion after PARPi in Patient 20

Figure 48: Genetic and clinical features in Patient 13, a *BRCA2* mutation carrier with breast cancer

Figure 49: Genetic and clinical features in Patient 6, a *BRCA1* mutation carrier with breast cancer

LIST OF ABBREVIATIONS

AAD, alternative allele depth (also Alt AD)
AAF, alternative allele fraction (also Alt AF)
BAM, binary alignment map (file format for aligned DNA/RNA sequences)
BRCA1, Breast Cancer Type 1 Susceptibility Protein (encoded by *BRCA1* gene)
BRCA2, Breast Cancer Type 2 Susceptibility Protein (encoded by *BRCA2* gene)
CGC, Cancer Gene Census
CNV, copy number variation
COSMIC, Catalog of Somatic Mutations in Cancer
DSB, DNA double strand break
ER+, estrogen receptor positive
ER-, estrogen receptor negative
FDR, false discovery rate
FFPE, formalin-fixed paraffin-embedded
gBRCA1/2, group of patients with a germline mutation in *BRCA1/2*
GO, gene ontology
GoF, gain of function (mutation)
GSEA, gene set enrichment analysis
HER2, Human epidermal growth factor receptor 2
HR, homologous recombination
HR-WT, group of patients without germline or somatic loss of any HR genes
HRD, homologous recombination deficiency
IGV, Integrative Genomics Viewer
IHC, immunohistochemistry
INDEL, small insertion/deletion
IRB, institutional review board
LoF, loss of function (mutation)
LOH, loss of heterozygosity (refers here to *BRCA1* or *BRCA2* zygosity)
LncRNA, long non-coding RNA
MAF, mutation annotation format (file format for mutations)
NGS, next generation sequencing
NHEJ, non-homologous end joining
NMD, nonsense mediated decay
nonLOH, absence of LOH (refers here to retention of one *BRCA1/2* wild type allele)
PARPi, poly(ADP-ribose) polymerase inhibitor
PARP1, poly(ADP-ribose) polymerase 1 (encoded by *PARP1* gene)
PR, progesterone receptor
RIP-seq, RNA immunoprecipitation sequencing
RNA-seq, RNA sequencing
SNV, single nucleotide variant
SNP, single nucleotide polymorphism
TCGA, The Cancer Genome Atlas program
TMB, tumor mutational burden
TNBC, triple-negative breast cancer (ER-/PR-/HER2-)
UTR, untranslated region (refers here to 3' UTR or 5' UTR of a gene)
WES, whole exome sequencing

CHAPTER 1: INTRODUCTION

BRCA1/2 Proteins in DNA Repair

Human cells have numerous mechanisms with which to repair damaged DNA. One common type of DNA lesion is the DNA double strand break (DSB), which is considered especially dangerous because it affects both strands. DSBs can be caused by normal DNA replication processes, ionizing radiation, and other genotoxic environmental or cell-intrinsic factors^{1,2}. If left unrepaired, DSBs can cause mutagenesis in key genes or mediate widespread genomic instability. Two main pathways exist to fix DSBs in eukaryotes: homologous recombination (HR) and non-homologous end joining (NHEJ)¹⁻³. HR is considered the more conservative and accurate route to DSB repair, because it utilizes an unaltered sister chromatid (containing a homologous DNA sequence) as a template strand for repairing the DNA lesion¹⁻⁴. HR can only occur during S or G2 phases of the cell cycle (when a sister chromatid is present), and its main functions are to repair stalled replication forks in addition to DSBs¹⁻⁴. This process involves many protein components, which are reviewed extensively elsewhere¹⁻³. Conversely, NHEJ is employed when a sister chromatid is unavailable; its machinery performs blunt end ligation of DNA fragments on either side of the break^{1,2,4}. NHEJ does not employ a template strand to ensure fidelity of the repaired DNA sequence; therefore, this mechanism is more error-prone than HR¹⁻⁴. Human cells use either HR or NHEJ to repair DSBs depending on stage of the cell cycle and availability of key proteins for each pathway³.

BRCA1 and BRCA2 proteins have distinct but critical roles in HR. BRCA1 largely works upstream of BRCA2, performing a wide range of direct and indirect functions to facilitate HR^{2,3}. BRCA1 is best known for two main roles. First, it promotes 5' to 3' end

resection of damaged DNA, thereby creating 3' overhangs that are required for HR³. Second, BRCA1 works with BRCA2 and other mediators (PALB2, etc.) to load RAD51 (DNA recombinase) onto resected ssDNA¹⁻³. In addition to these functions, BRCA1 complexes with a large repertoire of other proteins to sense DNA damage and mediate its repair³. While BRCA1 has well-studied roles in HR, the protein also functions in signal transduction, transcription, histone ubiquitination, and chromatin remodeling²⁻⁴. As mentioned above, BRCA2 is best known for a singular role: loading RAD51 onto ssDNA lesions¹⁻³. This process mediates the formation of RAD51 presynaptic filaments required for strand invasion, a key step in HR¹⁻³.

BRCA1/2 Mutations and Cancer Presentation

As a result of their roles in HR, BRCA1 and BRCA2 are well-characterized tumor suppressors¹⁻⁴. Both *BRCA1* and *BRCA2* genes were discovered by genomic linkage analyses and named based on their connection to hereditary **BR**east **CA**ncer⁵⁻⁸. *BRCA1/2* mutations are considered highly penetrant, based on apparent autosomal dominant inheritance patterns^{9,10}. Pathogenic *BRCA1/2* alleles are especially prevalent in ethnic groups with well-characterized founder mutations, including Ashkenazi Jews^{10,11}. In particular, carriers of heterozygous *BRCA1/2* mutations are highly susceptible to breast and ovarian carcinomas^{1,10}. *BRCA1* mutation carriers have (on average) a 57% chance of breast cancer and a 40% chance of ovarian cancer by age 70⁹. Similarly, *BRCA2* mutation carriers have a 49% chance of breast cancer and an 18% chance of ovarian cancer by age 70⁹. *BRCA1/2* mutation carriers are also predisposed to pancreatic and prostate cancers, although these were not a primary focus of this project^{4,10}.

BRCA1/2 mutation carriers also have specific disease presentation. *BRCA1/2* mutation carriers are predisposed to early onset cancers, sometimes developing several distinct primary tumors in their lifetimes^{1,10}. *BRCA1* mutation carriers are more likely to develop aggressive ER-/PR-/HER2- tumors (triple-negative breast cancer, or TNBC), for which treatment options are limited^{10,11}. In contrast, *BRCA2* mutation carriers tend to develop ER+ breast tumors and have a relatively later age of breast cancer onset^{1,12}. Relative to unselected breast cancer cohorts, *BRCA2* germline mutations also manifest in a high prevalence of male breast cancer¹⁰. Both *BRCA1/2* mutation carriers tend to develop serous ovarian carcinomas, although *BRCA1* mutation carriers develop them slightly earlier^{13,14}. Due to a lack of effective screening options, these ovarian carcinomas tend to be diagnosed as late-stage, high-grade malignancies^{10,13}. Because of their predisposition to cancer, *BRCA1/2* mutation carriers receive genetic counseling, enhanced screening, and prophylactic surgeries where indicated^{10,11}.

BRCA1/2 mutation-associated tumors typically receive the same frontline therapies as sporadic tumors of the same type. *BRCA1/2* mutation carriers with breast cancer are treated with chemotherapy, radiation, and surgery, which are informed by tumor characteristics and receptor status^{15,16}. *BRCA1/2* mutation-associated ovarian cancers are treated with primary debulking surgery and platinum/taxane chemotherapy¹³. For late-stage ovarian and breast cancer, poly(ADP-ribose) polymerase inhibitors (PARPi; Olaparib, Rucaparib, Niraparib, or Talazoparib) may be used as well^{15–21}.

BRCA1/2 mutation-associated tumors respond well to platinum-based chemotherapy and PARPi, as each type of therapy exploits their HR deficiency. Briefly, platinum (cisplatin, carboplatin) form protein-DNA adducts and DNA intra- and interstrand crosslinks, each of which mediate replication fork stalling and collapse^{4,22–24}. PARPi have

many effects, but their synthetic lethality stems from trapping PARP1/2 complexes on DNA, which also causes replication fork stalling and collapse^{25–27}. In both scenarios, widespread replication fork collapse causes DSBs that BRCA1/2-deficient cells cannot repair due to their HR deficiency^{24,26–28}. Ultimately, widespread DNA damage accumulates and causes apoptosis in these cells^{24,26–28}.

BRCA1/2 Loss of Heterozygosity

BRCA1/2 mutation-associated tumors develop as a result of DNA repair defects, which start with a heterozygous pathogenic *BRCA1/2* mutation. These cells ultimately accrue mutations and genomic rearrangements, which are thought to facilitate the development of cancer^{1,4}. Breast and ovarian epithelia are especially vulnerable to this process, and a recent study demonstrated that tumor suppression by *BRCA1/2* varies by cell lineage²⁹. As a hypothesis for *BRCA1/2* mutation-associated tissue tropism, others have suggested that hormonally-driven growth signals generate genotoxic reactive oxygen species in breast and ovarian epithelia¹. The relatively high degree of genotoxic stress in these tissues causes them to rely more heavily on HR¹.

BRCA1/2 mutation-associated tumors vary in their HR deficiency. In a typical *BRCA1* mutation carrier, normal cells contain one wild-type *BRCA1* allele and one mutant *BRCA1* allele. Typically the *BRCA1* mutation is inherited from a parent, although de novo germline *BRCA1* mutations have also been observed³⁰. Over the course of the patient's lifetime, some cells may lose the wild-type *BRCA1* allele due to somatic mechanisms (most commonly copy number losses)³¹. These cells are considered to have undergone allele-specific loss of heterozygosity, or LOH. The same occurs in *BRCA2* mutation carriers, upon loss of their wild-type *BRCA2* allele. LOH is thought to be an early event in most *BRCA1* and *BRCA2* mutation-associated tumors, as it has

been observed in normal tissue from *BRCA1/2* mutation carriers^{32,33}. This sequence of events follows Knudson's "two-hit hypothesis," the classical model by which biallelic loss of a tumor suppressor gene causes cancer³⁴.

In line with Knudson's two-hit hypothesis, it was long assumed that all *BRCA1/2* mutation-associated tumors underwent LOH^{35,36}. However, our lab recently demonstrated that a substantial minority of tumors do not undergo LOH, and instead retain their wild-type *BRCA1/2* allele³¹. These "nonLOH" tumors arise from *BRCA1/2* haploinsufficiency, rather than biallelic loss. As mentioned above, tumors deficient for *BRCA1/2* are correspondingly deficient for homologous recombination and sensitive to platinum and PARPi³⁷⁻⁴⁴. In contrast, patients without LOH may have primary resistance to platinum and PARPi due to their tumors' relative *BRCA1/2* (and HR) proficiency³¹. In support of this hypothesis, *BRCA1/2* mutation carriers with nonLOH, platinum-treated ovarian tumors had worse overall survival than those whose tumors had LOH³¹. Further, absence of *BRCA1* LOH has been observed in residual ovarian tumor following neoadjuvant platinum treatment, suggesting that this therapy may select for nonLOH clones⁴⁵⁻⁴⁷.

Acquired Therapeutic Resistance in BRCA1/2 Mutation-associated Tumors

BRCA1/2 mutation carriers are often diagnosed with high-grade serous ovarian carcinomas and hormone receptor-negative breast cancers, both of which are known to recur within a few years of diagnosis^{13,48}. While these patients initially respond to frontline therapies (including platinum), many of their tumors do return in this timeframe^{13,48}. Recurrent ovarian cancer and metastatic breast cancer are usually associated with acquired therapeutic resistance; as such, they are considered incurable. Breast tumors may recur locally (to nearby lymph nodes, breast, or the chest wall), or

distantly (to distant lymph nodes, lung, liver, brain, and bone)⁴⁸. Conversely, most ovarian cancer recurrences are local intra-abdominal recurrences, often with multi-organ involvement¹³. In general, cancer recurrence and metastasis are common causes of death in *BRCA1/2* mutation carriers, which is why we focused the study on these processes.

In *BRCA1/2* mutation carriers, the most common identified resistance mechanism to platinum-based chemotherapies and PARPi are reversion mutations in the mutant *BRCA1/2* allele^{49–57}. *BRCA1/2* reversions are defined as somatic mutations that “reverse” or correct the germline mutation (typically deleting the area around it), to restore *BRCA1/2* expression^{51,58}. Reversion events have been observed in several case studies of patients with *BRCA1/2* truncating mutations. However, the percentage of patients for which they constitute the mechanism of resistance is unknown, as most studies include very few (3-8) pairs of primary and recurrent tumors^{49–57}. Since reversion mutations are readily detectable in blood, prior studies may also be biased by the ease of collecting blood relative to tumor specimens. Moreover, truncating mutations (correctable by reversion) only comprise 77% of the germline *BRCA1/2* mutations found in carriers with breast or ovarian cancer⁵⁹. A recent meta-analysis found that the potential for *BRCA1/2* reversion varies by gene, mutation type, and mutation location within each gene⁶⁰. Lastly, *BRCA1/2* reversions have been identified following platinum and/or PARPi, but neither of these are frontline therapies for breast cancer^{16,61}. Therefore, it remains to be seen what role *BRCA1/2* reversions play in unselected, larger cohorts of recurrent *BRCA1/2* mutation-associated tumors after conventional treatment.

A number of studies have proposed other mechanisms of therapeutic resistance to platinum and PARPi in *BRCA1/2* mutation carriers. In the setting of *BRCA1* deficiency,

expression of hypomorphic *BRCA1* alleles can confer some HR capacity, as well as concomitant platinum and PARPi resistance^{62–65}. PARPi resistance may also arise from restoration of BRCA1-independent HR, which occurs via loss of TP53BP1 or MAD2L2/REV7^{66–69}. TP53BP1 normally functions to promote NHEJ and inhibit end resection, an early step in HR⁶⁷. BRCA1 normally functions to relieve this inhibition, enabling end resection and HR to proceed⁶⁷. Somatic loss of TP53BP1 renders this function of BRCA1 dispensable, thereby mediating end resection and PARPi resistance in BRCA1-deficient backgrounds^{67–69}. Similarly, MAD2L2 is thought to be a downstream effector of TP53BP1, and its loss also enables BRCA1-independent end resection and PARPi resistance⁶⁶. Of note, TP53BP1 and MAD2L2 losses appear to mediate resistance to PARPi, but not platinum, *in vitro* and *in vivo*²⁸. Neither mechanism has been reported in PARPi-resistant *BRCA1* mutation carriers.

TP53BP1 and MAD2L2 losses do not mediate platinum or PARPi resistance in BRCA2-deficient backgrounds^{66–69}. Instead, drug resistance may be conferred to BRCA2-deficient cells by loss of CHD4, a nucleosome remodeler⁷⁰. CHD4 normally inhibits translesion synthesis, an error-prone alternative DNA repair mechanism⁷⁰. Its loss permits translesion synthesis to proceed, an effect that facilitates survival of BRCA2-deficient cells treated with platinum or PARPi⁷⁰. Further, low *CHD4* expression has been correlated with reduced overall survival in *BRCA2* mutation-associated ovarian cancer⁷⁰. This effect was not observed for BRCA1-deficient backgrounds⁷⁰.

While the effects of TP53BP1, MAD2L2, and CHD4 loss vary by germline mutation, other mechanisms of platinum and PARPi resistance appear to be gene-agnostic. Overexpression of the ABC drug transporter MDR1 has been proposed as a mechanism of chemoresistance in *BRCA1/2* mutation-associated cancer, based on observations from

both breast and ovarian tumors^{51,71,72}. In each of these reports, MDR1 hyperactivity was mediated by the pump-encoding gene (*ABCB1*) fusing to an active promoter^{51,71}. Several types of PARPi are also predicted to be MDR1 substrates, adding further support for MDR1-mediated therapy resistance^{26,57}. Lastly, PARPi resistance may also be mediated by PARP1 loss, since PARPi rely on PARP1 for their synthetic lethal effect^{23,25,57,73}. For the same reason, PARP1 loss is not thought to mediate platinum resistance²³. Importantly, out of all resistance mechanisms described above, only *BRCA1/2* reversions and MDR1 expression have been observed in human *BRCA1/2* mutation-associated breast and ovarian tumors^{51,60,71}.

Remaining Gaps in Knowledge

In this study, we aimed to address existing gaps in our understanding of *BRCA1/2* mutation-associated cancers. One such gap is the lack of RNA sequencing (RNA-seq) in prior studies, since RNA-seq was historically infeasible in formalin-fixed paraffin-embedded (FFPE) samples. One recent study found that three recurrent *BRCA1/2* mutation-associated breast tumors differed (from matched primary tumors) in gene expression related to TNF α signaling, epithelial to mesenchymal transition (EMT), and stemness⁷¹. Another study in mouse models of TNBC identified *BRAF*, *RAF1*, and *FGFR2* fusion products, which increased oncogenic MAPK and PI3K signaling⁷⁴. While these studies suggest that differential gene expression contributes to *BRCA1/2* mutation-associated tumorigenesis, no one has assessed transcriptomic differences between primary and recurrent tumors on a broad scale. In this study, we sought to fill this knowledge gap and use transcriptomics to assess novel therapeutic vulnerabilities in recurrences.

This project represents the largest comprehensive analysis of paired primary and recurrent *BRCA1/2* mutation-associated tumors to date. Prior studies are limited by sample size (3-8 patients, compared to our 27), a single sequencing modality (compared to our multi-omic analysis), and a lack of matched tumor/germline samples (included here)^{49–57}. Importantly, the tumors collected for this study also reflect the standards of care for breast and ovarian cancer in *BRCA1/2* mutation carriers at the time of sample collection. Existing studies in the field often focus on platinum and PARPi resistance, as well as *BRCA1/2* reversion mutations known to mediate each. However, neither platinum nor PARPi are current frontline therapies for *BRCA1/2* mutation-associated breast cancer, and PARPi are only used in maintenance and metastatic settings for *BRCA1/2* mutation-associated ovarian cancer^{13,15–21,61}. For this reason, we posit that platinum and PARPi resistance does not cause initial disease recurrence in many *BRCA1/2* mutation carriers, particularly those with breast cancer.

While we looked for evidence of drug resistance mediated by the mechanisms described above, we focused more on identifying novel mechanisms of recurrence after conventional frontline treatments. Our aim was to identify genetic and transcriptomic events that underlie acquired therapeutic resistance in these patients. We hypothesized that differences between paired tumors would indicate general mechanisms of tumor evolution towards therapeutic resistance in HR-deficient cancers. Further, we hypothesized that multi-omic features shared by primary and recurrent tumors could be just as informative. In theory, such features could inform personalized medicine approaches for cancer recurrence, tailored to the genetic landscape of a patient's primary tumor. Overall, we hope this study improves our understanding of *BRCA1/2*

mutation-associated tumor biology, specifically to improve outcomes for patients with late-stage disease.

CHAPTER 2: PATIENT COHORT AND MULTI-OMIC SEQUENCING

Introduction

Prior studies of *BRCA1/2* mutation-associated cancers have been limited by low numbers of paired primary and recurrent tumors, making it difficult to know whether genetic features of recurrences are truly “recurrence-specific” drivers of therapeutic resistance. Some studies also lack matched germline DNA samples, making it difficult to tell whether mutations are truly somatic. Further, these studies have focused heavily on patients pre- and post-treatment with PARPi. While PARPi are approved for late-stage disease in *BRCA1/2* mutation-associated breast and ovarian cancer, they are not currently employed as frontline treatments for either. As a result, there exists a considerable knowledge gap as to how *BRCA1/2* mutation-associated tumors evolve after conventional frontline therapies. Lastly, prior studies have rarely included a transcriptomics component, so changes in gene expression or targetable gene fusion events have not been explored in tumor pairs. We performed multi-omic sequencing on the patient cohort described below, with the goal of addressing long-standing questions about *BRCA1/2* mutation-associated cancers.

Methods

Acquisition of tumor and germline specimens from *BRCA1/2* mutation carriers

Patients gave informed consent for research use of germline DNA, tumor specimens, and clinical data, under a University of Pennsylvania IRB approved protocol. Eligible patients met the following criteria: (1) diagnosis of breast or ovarian cancer; (2) positive genetic test result for pathogenic germline mutation in *BRCA1* or *BRCA2* from a Clinical Laboratory Improvement Amendments (CLIA)-approved laboratory; (3) available

archived germline DNA from blood or saliva; (4) available formalin-fixed paraffin-embedded (FFPE) blocks from the primary tumor; and (5) available FFPE blocks from at least one recurrence matching the primary tumor. Matched recurrent tumor samples were identified using manual review of patient charts and pathology reports, supervised by Dr. Susan M. Domchek and in collaboration with the Bassett Center for BRCA. In particular, we took care to match each primary tumor with its appropriate recurrence(s), which was especially critical given that *BRCA1/2* mutation carriers can develop distinct primary tumors over the course of their lifetimes⁷⁵. All clinical phenotypic data were collected through manual chart review. Patients 1-27 met all study criteria and comprised the primary/recurrent patient cohort. An additional four patients (Patients 28-31) lacked an available primary tumor sample (meeting all but criterion 4) and were included only in groupwise RNA-seq analyses.

Acquisition of normal tissue from *BRCA1/2* mutation carriers without cancer

We also identified 12 *BRCA1/2* mutation carriers as a control group for RNA-seq analyses. Eligible patients met the following criteria: (1) positive genetic test result for pathogenic germline mutation in *BRCA1* or *BRCA2* from a CLIA-approved laboratory; (2) no prior history of cancer or cancer treatment; and (3) available FFPE blocks from prophylactic mastectomy or salpingo-oophorectomy surgeries. We extracted and sequenced RNA from normal breast and fallopian tube samples from these 12 patients. Of note, fallopian tube samples were collected as controls for ovarian cancer, since fallopian tube (not ovary) is thought to be the tissue of origin⁷⁶.

Pathological review of FFPE specimens

FFPE tumors were collected and sectioned by the Tumor Tissue and Biospecimen Bank at the University of Pennsylvania, then stained using hematoxylin and eosin. For tumor samples, staining was reviewed and marked by Dr. Anupma Nayak (pathologist) to identify sections with $\geq 50\%$ invasive tumor. Normal tissue samples (RNA-seq controls) were reviewed by Dr. Nayak to confirm the absence of tumor. DNA and RNA were extracted from FFPE sections or rolls of 5-10 μ m thickness.

DNA extraction for whole exome and targeted sequencing

DNA was extracted from FFPE samples using Deparaffinization Solution (Qiagen) and Gentra PureGene Tissue Kits (Qiagen). For a minority of tumors, DNA was extracted from FFPE using standard laboratory deparaffinization, proteinase K digestion, and ethanol precipitation. Germline DNA was extracted from whole blood using sucrose-based lysis of erythrocytes, followed by proteinase K digestion and ethanol precipitation from leukocytes. Germline DNA was extracted from saliva samples using Oragene kits (DNA Genotek).

RNA extraction for RNA sequencing

Tumor RNA was extracted from the same FFPE tumor blocks from which DNA was extracted. Tumor and normal RNA samples were extracted from FFPE with the RNeasy FFPE kit (Qiagen). To preserve RNA integrity, RNA samples were extracted within one week of sectioning, handled using RNaseZap and barrier pipette tips, and stored at -80°C. RNA samples underwent a maximum of two freeze/thaw cycles.

Whole exome and targeted sequencing of tumor and germline DNA

All DNA was sheared for library preparation using a Covaris sonicator. Tumor DNA libraries were prepared using the NEBNext FFPE Repair mix and NEBNext Ultra II DNA library prep kit (New England Biolabs), per manufacturer's instructions. Germline DNA libraries were prepared using the NEBNext Ultra DNA library prep kit (New England Biolabs), per manufacturer's instructions. DNA libraries were pooled and hybridized using SureSelect Target Enrichment System for Illumina Multiplex Sequencing (Agilent) and associated protocols. For whole exome sequencing (WES), tumor and germline libraries were hybridized to SureSelect All Exon v5, SureSelect All Exon v6+COSMIC, and SureSelect All Exon v7 captures (Agilent). For targeted sequencing, tumor and germline libraries were hybridized to a custom capture, which largely utilized baits from the SureSelect All Exon platform (Agilent; see **Table 2** for gene list). We quantified concentrations of DNA samples, libraries, and hybridization pools using a Qubit (ThermoFisher). We assessed quality (fragment size) of DNA libraries and hybridization pools using a BioAnalyzer 2100 (Agilent). WES was performed using an Illumina HiSeq 4000 and targeted sequencing was performed using an Illumina NovaSeq 6000. All sequencing was performed with 150 paired-end reads by the University of Pennsylvania Next Generation Sequencing Core. We aimed to achieve a mean depth of 40x in germline and 100x in tumor by WES; any samples with depth < 20x were re-sequenced. For targeted sequencing, we aimed to achieve a mean depth of 200-300x for both tumor and germline DNA.

RNA sequencing of tumor and normal RNA

We used the TruSeq RNA Exome platform (Illumina) for library preparation and hybridization capture, per manufacturer protocols. Sequencing was performed with 150 paired-end reads using an Illumina HiSeq 4000 at the University of Pennsylvania Next

Generation Sequencing Core. We aimed to sequence at least 20 million reads per sample.

Sequencing alignment for DNA and RNA sequencing

Fastq files from whole exome and targeted sequencing were aligned to the hg19 build of the human genome using the Burrows-Wheeler Aligner (BWA v.0.7.17-r1188)⁷⁷. The resulting binary alignment map (bam) files were processed according to Genome Analysis Toolkit (GATK v3.7) best practices (picardtools v2.20.7)⁷⁸. Fastq files from RNA-seq were aligned using STAR aligner (v2.7.2a) to gene annotations from the GENCODE Human Release 19 reference assembly (available at https://www.gencodegenes.org/human/release_19.html)⁷⁹.

Results

We collected a cohort of 27 breast and ovarian cancer patients with pathogenic germline mutations in *BRCA1* or *BRCA2* (see **Methods**). Patients were identified, and their samples acquired, in collaboration with the Bassett Center for BRCA at the University of Pennsylvania. Tumor DNA, tumor RNA, and normal tissue RNA were all extracted from FFPE blocks. Germline DNA was extracted from saliva or blood samples. We also collected a panel of normal control tissue for RNA sequencing controls, which consisted of six normal breast and six normal fallopian tube specimens from prophylactic surgeries in *BRCA1/2* mutation carriers (see **Methods**).

Cohort selection and clinical characteristics

We studied 13 matched primary and recurrent breast cancers from nine *BRCA1* and four *BRCA2* pathogenic variant carriers (**Table 1**). Nine of 13 primary breast tumors

were negative for estrogen receptor expression (ER-); these were predominantly triple-negative breast cancer specimens (TNBC; 17/19 tumors) with concordant ER-recurrences. We evaluated matched primary and recurrent ovarian tumors from 14 women, including 10 *BRCA1* and four *BRCA2* pathogenic variant carriers. Most (9/14) primary ovarian tumors were diagnosed at stage III or IV. With the exception of one male *BRCA2* mutation carrier with breast cancer, all patients were female. Recurrences were both local and distant, and generally represented the most common sites of recurrence for each cancer type (**Fig. 1**).

As treatment for their primary tumor, all patients received chemotherapy, with 8/27 (29%) also receiving hormonal therapy and 10/27 (37%) receiving radiation (**Table 1**). Seventeen (63%) patients received platinum-based chemotherapy, which included three breast cancer and all (14) ovarian cancer patients¹³. Four ovarian cancer patients and one breast cancer patient received PARPi prior to the sampled recurrence. Additional clinical and sequencing details for each tumor are documented in **Supplementary File 1**.

Multi-omic Sequencing by Sample

All tumor sequencing by sample is depicted in **Figure 2**. We collected a total of 67 tumors (27 primary and 40 recurrent) for whole exome sequencing. For each of the 27 patients in the cohort, whole exome sequencing was performed on the primary tumor and at least one recurrence. Additionally, we re-sequenced DNA from 44 of 67 tumors at high depth using a custom targeted capture. Whole exome and targeted sequencing were conducted on matched germline DNA for each tumor in parallel. Next, we sequenced RNA from 50 of 67 tumors in the WES cohort, along with four additional

recurrences that lacked a matched primary tumor. In parallel with tumor RNA-seq, we sequenced RNA from 12 normal breast and fallopian tube samples.

WES achieved a mean depth of 98x in tumor samples (median depth 87x) and 93x in germline samples (median depth 94x). Targeted sequencing achieved a mean depth of 332x in tumor samples (median depth 304x) and 317x in germline samples (median depth 316x). RNA sequencing achieved a mean depth of 67.6 million reads per sample (median 55.6 million reads) for tumor and normal specimens.

The main outcomes from WES were somatic variants and allele-specific copy number variation (**Fig. 3**). Results from mutations and copy number variation (with related downstream analyses) are discussed in chapters 3 and 4, respectively. Targeted sequencing was used for more sensitive detection of somatic variants in select genes (**Table 2**) and is also discussed in chapter 3. The main outcomes from RNA-seq were detection of gene fusions, sample clustering, differential gene expression, and related downstream analyses (**Fig. 4**). Each of these is discussed in chapter 5. Given the high depth of RNA-seq coverage, we also assessed differential transcript usage (**Fig. 4**), which is discussed in chapter 6. Results from multi-omic sequencing are integrated to trace tumor evolution in three patients in chapter 7.

Table 1: Patient cohort and tumor characteristics

		Breast	Ovarian	Total
Patients		13	14	27
Tumors				
	Primary	13	14	27
	Recurrent	18	22	40
	Total	31	36	67
Recurrences Collected by Patient				
	1	9	10	19
	>1	4	4	8
Germline Mutation				
	<i>BRCA1</i>	9	10	19
	<i>BRCA2</i>	4	4	8
ER Status of Primary Tumor				
	ER+	4	N/A	4
	ER-	9	N/A	9
Stage at Diagnosis				
	I	2	1	3
	II	8	3	11
	III	2	8	10
	IV	0	1	1
	Unreported	1	1	2
Treatment Received^{a,b}				
	Platinums	3	14	17
	PARPi ^c	1	4	5
	Chemotherapy	13	14	27
	Hormonal Therapy	5	3	8
	Radiation	7	3	10
^a Treatments are limited to reflect the number of patients from whom a post-treatment recurrence was collected.				
^b Treatment groups are not mutually exclusive; most patients received multiple treatments.				
^c PARPi = Poly (ADP-ribose) polymerase inhibitor				

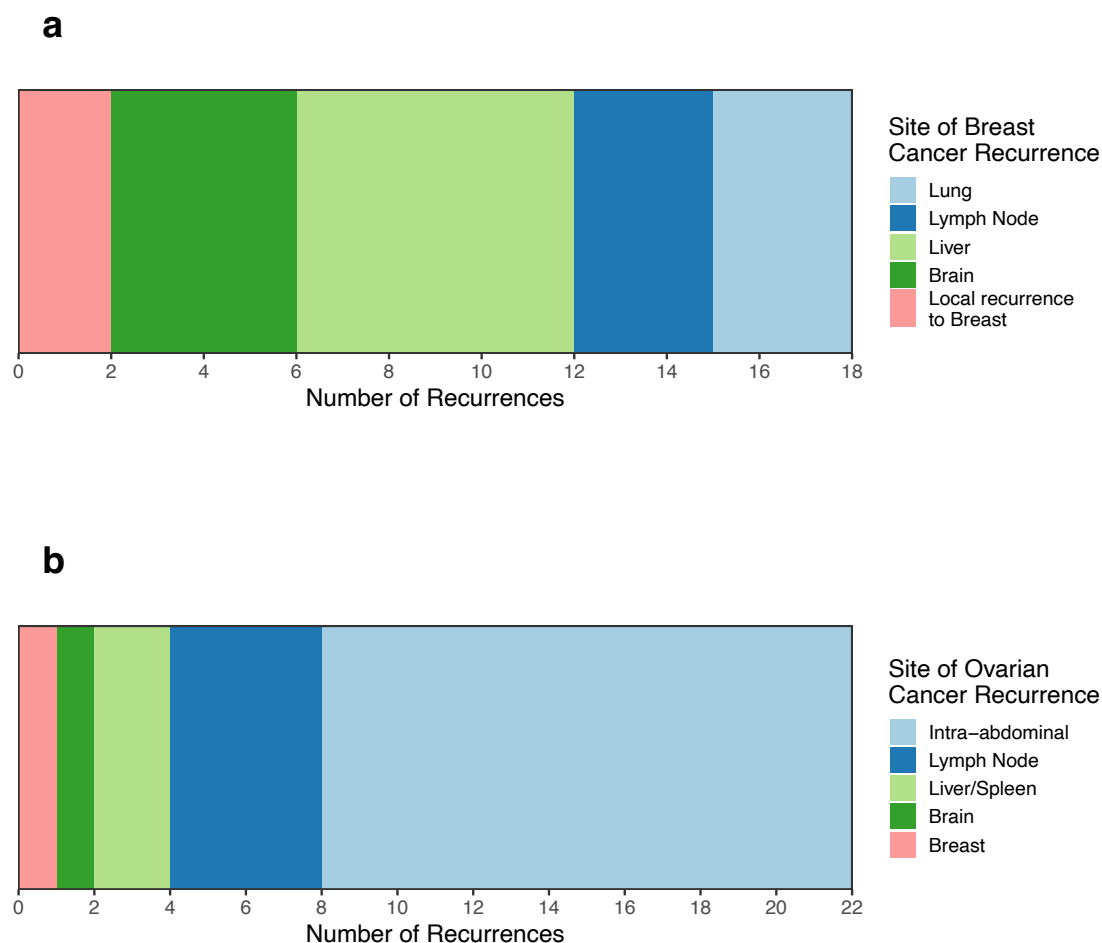


Figure 1: Distribution of recurrences collected by anatomical site. A. Anatomical locations of 18 breast cancer recurrences collected for multi-omic sequencing. B. Anatomical locations of 22 ovarian cancer recurrences collected for multi-omic sequencing. “Intra-abdominal” ovarian recurrences refer to those found in reproductive and gastrointestinal tracts, with many invading multiple organs therein. Lymph node recurrences in A and B were from both local and distant sites.

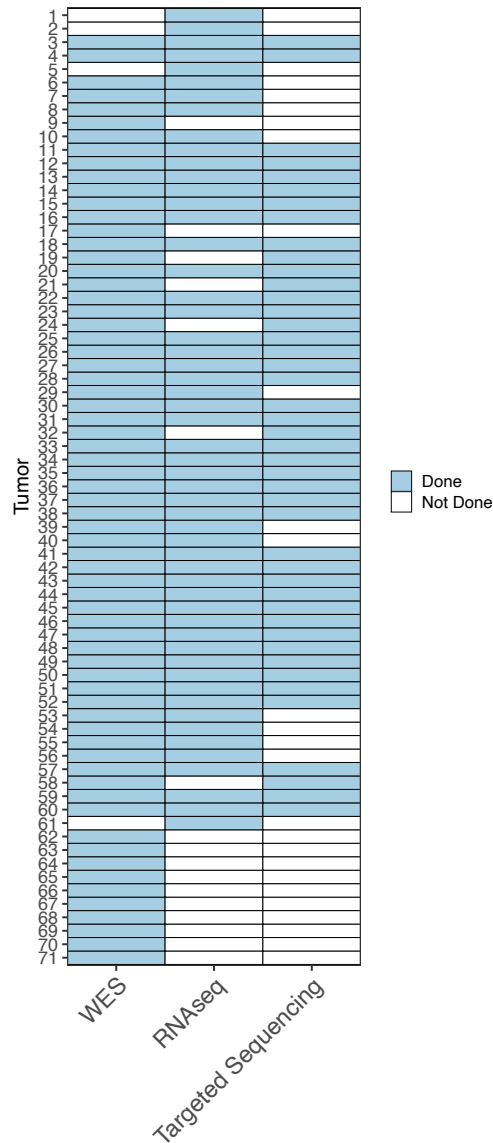


Figure 2: Multi-omic sequencing completed by tumor. Next-generation sequencing performed on DNA and/or RNA for each tumor in the cohort. Tumors 1, 2, 5, and 61 were unpaired ovarian recurrences from *BRCA1/2* mutation carriers, used solely for groupwise RNA-seq analyses. All 67 other tumors received whole exome sequencing (WES). Normal tissue controls for RNA-seq are not displayed. Tumor numbering is arbitrary; see figures in later chapters for tumor metadata displayed by patient.

Table 2: Genes included in targeted capture

Exons+UTR Coverage				Whole Gene Coverage	
ABL1	DICER1	MAP2K2	RAD51	ABCB1	TGFB2
AKT1	DIRAS3	MAP2K4	RAD51B	APC	TOP2B
AKT2	DLEC1	MAP3K4	RAD51C	ATXN3	TP53
AKT3	DPH1	MAPK1	RAD51D	BRCA1	TSC1
ALK	EIF5A2	MAPK3	RAD52	BRCA2	TSC2
AQR	EME1	MECOM	RAD54B	CCNE1	XRCC1
ARAF	EME2	MET	RAD54L	CDH1	
ARID1A	EPHA3	MLH1	RASSF1	CHD4	
ARL11	ERBB2	MPL	RFC1	CUL4B	
ATM	ERBB3	MRE11A	RNF111	EGFR	
ATR	ERBB4	MSH3	RP11-554I8.2	EP300	
AURKA	ERCC2	MSH4	RP3-510D11.2	FGFR2	
BAP1	ERCC3	MSH6	RPA1	FGFR3	
BARD1	ERCC4	MTOR	RPA2	FLT4	
BLM	ERCC6	MUS81	RPA3	FOXA1	
BRAF	ESR1	NBN	RPA4	GATA3	
BRIP1	FANCA	NF2	RPS6KA2	JAK1	
C11ORF80	FANCE	NOTCH2	RSF1	MAD2L2	
CBX2	FANCM	NOTCH3	SETMAR	MAP3K1	
CCND1	FBXW7	NRAS	SHFM1	MDM2	
CCND2	FGF1	OPCML	SLX4	MSH2	
CCND3	FGFR1	PALB2	SMAD2	MYC	
CDK12	FLT3	PAX8	SMARCA4	NCOR1	
CDK4	FOXL2	PEG3	SMARCB1	NF1	
CDK6	FRS2	PIK3R1	SMCHD1	NOTCH1	
CDKN1A	GNA11	PLAGL1	SPARC	PARP1	
CDKN1B	GNAQ	PMS2	SRC	PARP2	
CDKN1C	GNAS	POLD1	SSBP1	PDGFRA	
CDKN2A	GTF2H3	POLD2	STK11	PIK3CA	
CDKN2B	HRAS	POLD3	TERT	POLE	
CEBPA	IGF1R	POLD4	TOP2A	POLR2A	
CHEK1	JAK2	POLK	TOP3A	PRDM7	
CHEK2	JAK3	PPIE	TOP3B	PTEN	
CREBBP	JARID2	PPM1D	TP53BP1	RAD23B	
CRKL	KDR	PPP2R1A	TRIP13	RAF1	
CSF1R	KIT	PRDM9	WVVOX	RB1	
CTNNB1	KLHDC3	PRKCI	XRCC2	RBBP8	
CUL4A	KMT2C	PRKDC	XRCC3	RIC8A	
DAB2	KRAS	PRPF19	XRCC6	RIF1	
DDB1	LIG3	RAB25		SMAD4	
DDR2	MAP2K1	RAD50		TBX3	

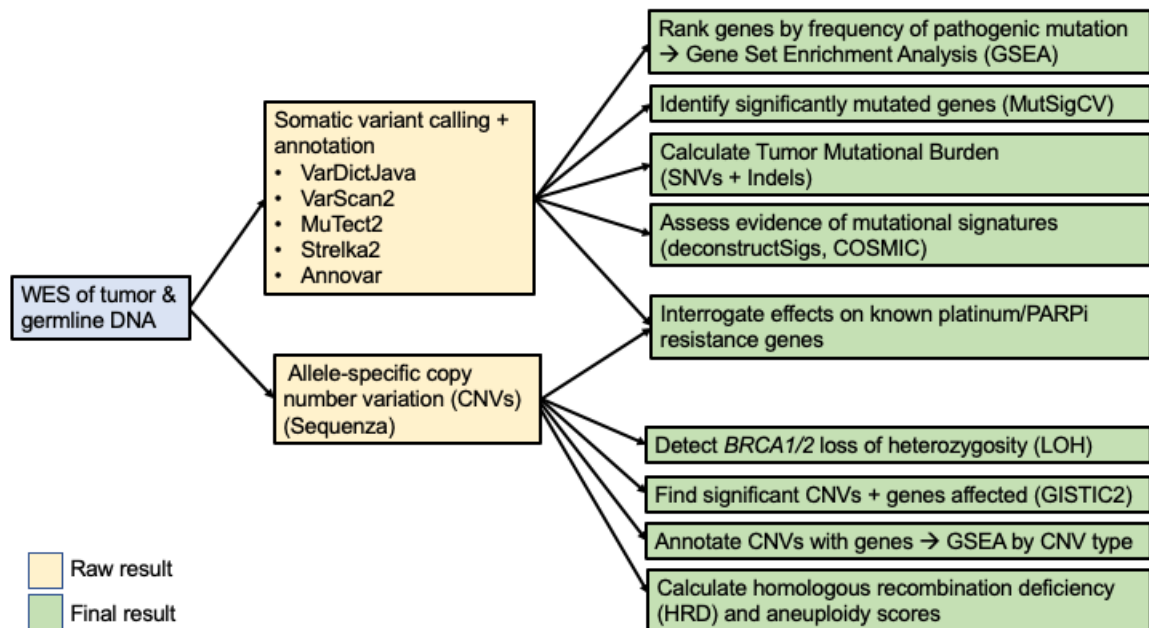


Figure 3: Overview of analyses from whole exome sequencing. Flowchart depicting the main outcomes from tumor and germline whole exome sequencing (WES). Green boxes indicate final results, which are reflected in subsequent figures. These results were drawn from somatic variant calls and allele-specific copy number, and are presented in chapters 3 and 4, respectively.

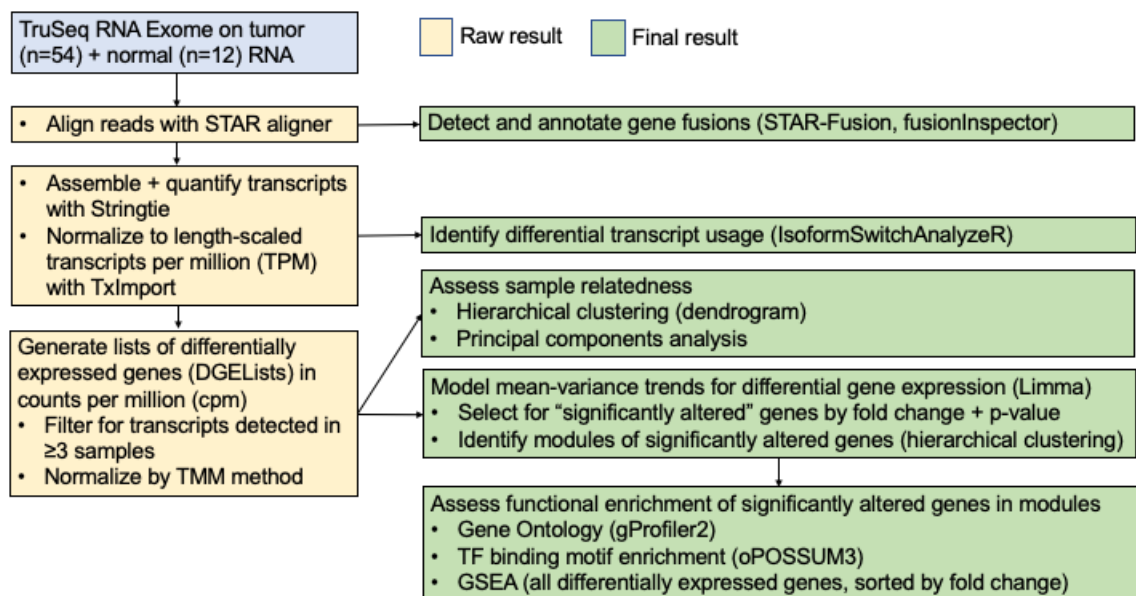


Figure 4: Overview of analyses from RNA sequencing. Flowchart depicting the main outcomes from tumor and normal RNA-seq. Green boxes list final results, which are reflected in subsequent figures. Differential transcript usage is discussed in chapter 6; all other results are discussed in chapter 5.

Discussion

The patients and tumors collected for this study reflect broader trends within *BRCA1/2* mutation-associated cancers. Breast tumors represented in the cohort were mostly triple-negative breast cancer (TNBC), in line with the high prevalence of TNBC in *BRCA1/2* mutation-associated breast tumors⁴⁸. Ovarian tumors tended to be high-grade and late-stage, as expected due to lack of effective screening methods for ovarian carcinomas in general¹³. Overall, the cohort was slightly skewed towards *BRCA1* mutation carriers (19/27 patients), which reflects the higher prevalence of breast and ovarian cancer in *BRCA1* compared to *BRCA2* mutation carriers⁹. Since PARPi are not

used as frontline therapies for either tumor type (and were less commonly used in general at the time of specimen collection), our sample set of post-PARPi recurrences was accordingly small. As platinum is the standard of care for ovarian cancer currently, we collected at least one post-platinum recurrence for each ovarian cancer patient⁸⁰. Ultimately, most recurrences were exposed to chemotherapy and radiation, reflecting common frontline treatment strategies.

Our cohort of patients and their samples had some limitations. We attempted to use the same DNA samples for whole exome and targeted sequencing, and RNA from the same FFPE tumor block for RNA sequencing, but it was not always possible due to sample availability. As a result, some multi-omic results from the same tumor might disagree due to tumor heterogeneity. Further, we only sequenced DNA or RNA from one FFPE block per sample, when an ideal approach would have been to sequence samples taken from multiple sites for larger tumor surgeries. Lastly, while breast tumors occasionally metastasize to bone, it is difficult to extract high-quality DNA from bone specimens. Therefore, our collection of breast tumors does not include any recurrences to bone, and we cannot make any conclusions about the applications of these findings to bone metastases. Overall, although our final cohort was relatively small, it represents the largest set of paired primary and recurrent *BRCA1/2* breast and ovarian tumors collected to date.

CHAPTER 3: ANALYSIS OF SOMATIC MUTATIONS

Introduction

BRCA1 and BRCA2 proteins are key nodes in the homologous recombination pathway, which is considered the most conservative and faithful route to DNA double strand break repair^{1,2,4–8,81,82}. In *BRCA1/2* mutation carriers, partial or total loss of BRCA1/2 forces cells to rely on error-prone DNA repair pathways, such as non-homologous end joining^{1,4}. This switch can cause cells to accumulate mutations and chromosomal abnormalities which mediate the transition to cancer^{1,4}. As in all cancers, accumulated mutations can be “drivers” which contribute to tumor initiation and progression, or “passengers” that reflect the increase in DNA lesions but are less biologically relevant themselves⁸³.

In this study, we used several orthogonal approaches to assess somatic variants in *BRCA1/2* mutation-associated breast and ovarian tumors. Broadly, our goal was to identify driver mutations and key pathways affected by mutations in this cohort. We also examined mutations in genes related to platinum and PARPi resistance. In general, we sought to understand how mutational profiles evolve over the course of tumor recurrence. We assessed mutational signatures, which are global patterns of single nucleotide variants (SNVs) known to be correlated with specific mutagenic processes⁸⁴. We also calculated tumor mutational burden (TMB), which represents the total count of SNVs and small insertions/deletions (Indels) in a given tumor⁸⁵. Lastly, a main goal of our variant calling was to identify somatic *BRCA1/2* reversion mutations. While reversion mutations are anecdotally reported in the literature, we wanted to establish their prevalence in an unbiased cohort not limited to post-platinum and post-PARPi recurrences.

To assess the landscape of somatic mutations in *BRCA1/2* mutation-associated tumors, we used WES of 67 paired primary and recurrent *BRCA1/2* mutation-associated tumors with matched germline samples. We also utilized high-depth targeted sequencing (>300x) from 44 tumors and 18 respective germline samples, which granted us increased sensitivity of variant detection in 209 genes of interest (**Table 2**).

Methods

Identification, filtering, and analysis of somatic variants

We used a union of MuTect2 (v4.1.8.1), Strelka2 (v2.9.2), VarDictJava (v1.5.1), and VarScan2 (v2.4.4) to call somatic variants from whole exome and targeted sequencing of matched tumor and germline DNA^{86–89}. VarDictJava and VarScan2 were used to call and filter germline variants^{86,89}. Each patient's germline variant calls were checked to ensure detection of the same pathogenic *BRCA1/2* variant reported in their genetic test results. All variants were annotated with ANNOVAR (October 2019 release)⁹⁰. We limited our analysis of mutational signatures to variants called by MuTect2 and with alternative allele depth of ≥ 10 reads. Mutational signatures were assessed separately for breast and ovarian tumors using deconstructSigs (v1.8)^{84,91}.

With the exception of mutational signatures, all downstream analyses were restricted to exonic variants that were called by and passed filtering for at least one variant caller and had alternative allele read depth (AAD) of \geq five reads. Tumor mutational burden (TMB) was calculated for each tumor using the following equation⁸⁵:

$$TMB = \frac{(n \text{ somatic, exonic nonsynonymous SNVs} + \text{indels})}{\text{size of WES capture (Mbp)}}$$

In order to exclude common single nucleotide polymorphisms (SNPs) from further analyses, we next removed somatic variants with frequency in any population (PopFreqMax) ≥ 0.01 ⁹⁰. Using the filtered variant set, we identified significantly-mutated genes within primary and recurrent tumor groups using MutSigCV (v1.3)⁸³.

Next, we identified gain of function (GoF) and loss of function (LoF) mutations in individual genes. LoF mutations were identified as those that met one of the following criteria: (1) frameshift mutation, (2) nonsense mutation, or (3) nonsynonymous SNV predicted to be pathogenic by REVEL score > 0.5 ⁹². GoF mutations were defined as those that met all of the following criteria: (1) not a LoF mutation; (2) documentation in Catalog of Somatic Mutations in Cancer (COSMIC v84, <https://cancer.sanger.ac.uk/cosmic>); (3) occurring in a Tier 1 oncogene as defined by the Cancer Gene Census (CGC, <https://cancer.sanger.ac.uk/census>); and (4) matching the mutation type determined to be oncogenic for that gene by CGC. All mutation filtering strategies are represented graphically in **Figure 5**.

Pathway analysis of somatic variants

Pathway analysis was performed on LoF variants using Gene Set Enrichment Analysis (GSEA) Preranked (v4.0.2 for Windows)⁹³. Briefly, genes were ranked based on frequency of LoF mutations in primary and recurrent tumor groups separately. A gene could be counted as mutated only once per patient within each tumor group. This approach was used to prevent skewing of results by patients with multiple recurrences or by individual tumors accruing distinct LoF mutations in the same gene. Ranked gene lists were used as input for GSEA Preranked with Hallmark gene sets (1000

permutations). Gene set enrichment was considered significant using a cutoff of FDR < 0.25.

Detection of *BRCA1/2* biallelic loss via somatic mutations

We checked whether *BRCA1/2* biallelic loss occurred via secondary somatic mutations as detected by whole exome or targeted sequencing. Specifically, we determined whether somatic LoF mutations in *BRCA1/2* were present at an alternative allele fraction (AAF) of ≥ 0.25 . This cutoff was imposed to select for pathogenic secondary mutations that were present in enough of the tumor bulk to facilitate loss of the wild-type *BRCA1/2* allele.

Detection of *BRCA1/2* reversion via somatic mutations

We assessed potential *BRCA1/2* reversions on a case-by-case basis. For each *BRCA1* mutation detected by whole exome or targeted sequencing in a *BRCA1* mutation carrier, we manually assessed whether the mutation would “reverse” the patient’s documented germline mutation. The same was done for *BRCA2* mutations in *BRCA2* mutation carriers. Somatic *BRCA1/2* variants were considered reversions if they met one of the following criteria: (1) missense mutation that corrected a germline missense mutation to the reference sequence; (2) missense mutation that altered a germline premature termination codon to encode an amino acid instead; (3) in-frame deletion around the site of a germline missense, nonsense, or frameshift mutation; or (4) frameshift mutation downstream of a germline frameshift mutation, which would restore the lost open reading frame. Functional effects of somatic *BRCA1/2* variants were also assessed using MutationTaster (<http://www.mutationtaster.org/>) as needed⁹⁴. All potential *BRCA1/2* reversion mutations were also manually inspected in respective binary

alignment map (bam) files using the Integrative Genomics Viewer (IGV; IGV_2.8.13 for Mac)⁹⁵.

Statistical analysis

Paired comparisons of tumor mutational burden were tested for significance with two-sided Wilcoxon signed rank tests ($\alpha=0.05$). For patients with multiple recurrences, one recurrence was chosen at random for pairwise comparisons (such that $n=27$ for each group).

Results

Global analysis of mutations from whole exome and targeted sequencing

We performed ensemble somatic variant calling with MuTect2, VarDict, VarScan2, and Strelka2 to identify putative loss of function (LoF) and gain of function (GoF) mutations within each dataset (**Fig. 5** and see **Methods**)^{86–89}. We assessed individual gene variants to identify shared oncogenes or tumor suppressors. The majority of the cohort (53/67 tumors; 80%) had LoF mutations in *TP53* (**Fig. 6; Fig. 7a, b**). Breast tumors without *TP53* mutations were mostly *BRCA2* mutation-associated and ER-positive (ER+) in origin (8/9 tumors from patients 10-13). Five ovarian tumors from three patients (17, 20, 25) also lacked *TP53* mutations. *KMT2C* was mutated in 21/67 tumors, 16 of which were ovarian tumors. *NOTCH1* mutations were identified exclusively in (six) breast tumors, whereas *NF1* was mutated in six breast and ovarian tumors. One *BRCA2* carrier had an ER+ primary breast tumor and two recurrences, each of which lacked *TP53* mutations but instead appeared to be driven by the well-characterized oncogenic *AKT1* E17K (GoF) mutation^{96,97}. Other GoF mutations were identified in *PIK3CA* and *TBX3* in one breast and one ovarian tumor, respectively. All

displayed somatic LoF and GoF mutations were deemed heterozygous by the respective variant callers.

Next, we assessed mutational signatures to identify the processes underlying somatic mutations in these tumors (**Fig. 7c**). With few exceptions, mutational signatures were dominated by Signature 3 (*BRCA1/2* inactivating mutations and defective HR) and Signatures 1A/B (aging)⁸⁴. Sixteen of 27 (60%) tumors had contributions from Signature 4 (smoking)⁸⁴. We also evaluated tumor mutational burden. In paired comparisons between primary and recurrent tumors, we found that tumor mutational burden does not increase over the course of tumor evolution (**Fig. 8, Supplementary File 2**).

To investigate the functional impact of somatic variants, we determined whether LoF variants accumulated in common pathways across the cohort. We ranked genes by prevalence of LoF variants within primary and recurrent tumor groups, then performed Gene Set Enrichment Analysis (GSEA) Preranked with Hallmark Gene Sets on ranked gene lists (**Fig. 9, Supplementary File 3**)⁹³. LoF variants from primary and recurrent tumors fell into seven shared gene sets, highlighting shared potential dysregulation of signals for DNA repair, G2M checkpoint, mitotic spindle, Wnt β -catenin signaling, and epithelial to mesenchymal transition (EMT). LoF variants from recurrences were exclusively enriched for sixteen gene sets, highlighting potential differences in metabolic processes (oxidative phosphorylation, glycolysis), apoptosis, TP53 signaling, and pathways regulating cell growth and identity (TGF β , Notch, MTORC1, KRAS).

To identify additional driver mutations in *BRCA1/2* mutation-associated tumors, we performed MutSigCV analysis on primary and recurrent tumor groups (**Fig. 10; Supplementary File 4**)⁸³. *TP53* and *BAGE2* were the only significantly mutated genes

within both tumor groups (all FDR $q < 0.05$). *BAGE2* encodes a long noncoding RNA (lncRNA) previously characterized as a tumor antigen in melanoma, which may represent a novel tumor antigen for *BRCA1/2* mutation-associated breast and ovarian tumors⁹⁸. This cohort of primary and recurrent tumors appears to be predominantly driven by the pathogenic variants in *TP53*. However, seven patients (four breast, three ovarian) developed tumors without *TP53* mutations, suggesting the presence of additional drivers (including *AKT1*) in this sample set.

Mutations in genes related to platinum and PARPi resistance

We interrogated mutations from whole exome and targeted sequencing in genes associated with platinum and PARPi resistance in the literature (*ABCB1*, *TP53BP1*, *MAD2L2/REV7*, *CHD4*, and *PARP1*)^{23,51,66–68,70,71,73,99}. First, we looked for *ABCB1* mutations that could facilitate chemoresistance via hyper-activation of the encoded ATP-binding cassette (ABC) drug transporter. We identified one such (potential) GoF mutation in the primary ovarian tumor from patient 17, a *BRCA1* mutation carrier. This mutation, *ABCB1* c.C988G (p.Q330E), occurred in the ABC transporter type I transmembrane domain of *ABCB1* (per InterPro) and is predicted benign per REVEL score. However, this tumor also had a predicted LoF mutation nearby in the same domain (*ABCB1* c.G986A (p.G329E), REVEL score = 0.922), so the functional relevance of the GoF mutation is unclear.

Next, we looked for loss of function mutations in DNA damage response genes *TP53BP1*, *PARP1*, *MAD2L2/REV7*, and *CHD4*. We identified one *TP53BP1* pathogenic frameshift mutation in a breast cancer recurrence from a *BRCA1* mutation carrier (Patient 8, *TP53BP1* c.5325_5346del, a 20bp deletion in exon 25). This *TP53BP1*

mutation was detected by both whole exome and targeted sequencing (AAF = 0.267 and AAF = 0.271, respectively). We also found one pathogenic nonsense mutation in *PARP1* in patient 20's second recurrence (*PARP1* c.C844T (p.R282X), AAF=0.071). No pathogenic mutations were found exceeding an AAF of 0.05 for *CHD4* or *MAD2L2/REV7*.

Analysis of *BRCA1/2* loss of heterozygosity by secondary mutations

Lastly, we sought to determine whether primary and recurrent tumors had undergone *BRCA1/2* loss of heterozygosity via secondary somatic mutations in these genes. We looked for somatic pathogenic variants in *BRCA1/2* from whole exome and targeted sequencing, identifying pathogenic *BRCA1/2* variants in 20 of 67 tumors. However, each of these somatic *BRCA1/2* mutations had low alternative allele fraction (AAF < 0.045), suggesting that they only affected a minority of cells. Therefore, we concluded that none of these mutations could have facilitated complete loss of the wild-type *BRCA1/2* allele.

Analysis of *BRCA1/2* reversion by secondary mutations

We did not identify any *BRCA1/2* reversion mutations in 23 post-platinum recurrences with whole exome and/or targeted sequencing. Six of these post-platinum recurrences were also exposed to PARPi; and within this post-platinum/PARPi group, we found the same *BRCA1* reversion in two tumors from one patient. This patient was a *BRCA1* mutation carrier with ovarian cancer (Patient 20). The tumors' reversion mutation was a somatic 15bp in-frame deletion around the site of her germline *BRCA1* nonsense mutation. The reversion and Patient 20's case are discussed further in chapter 7.

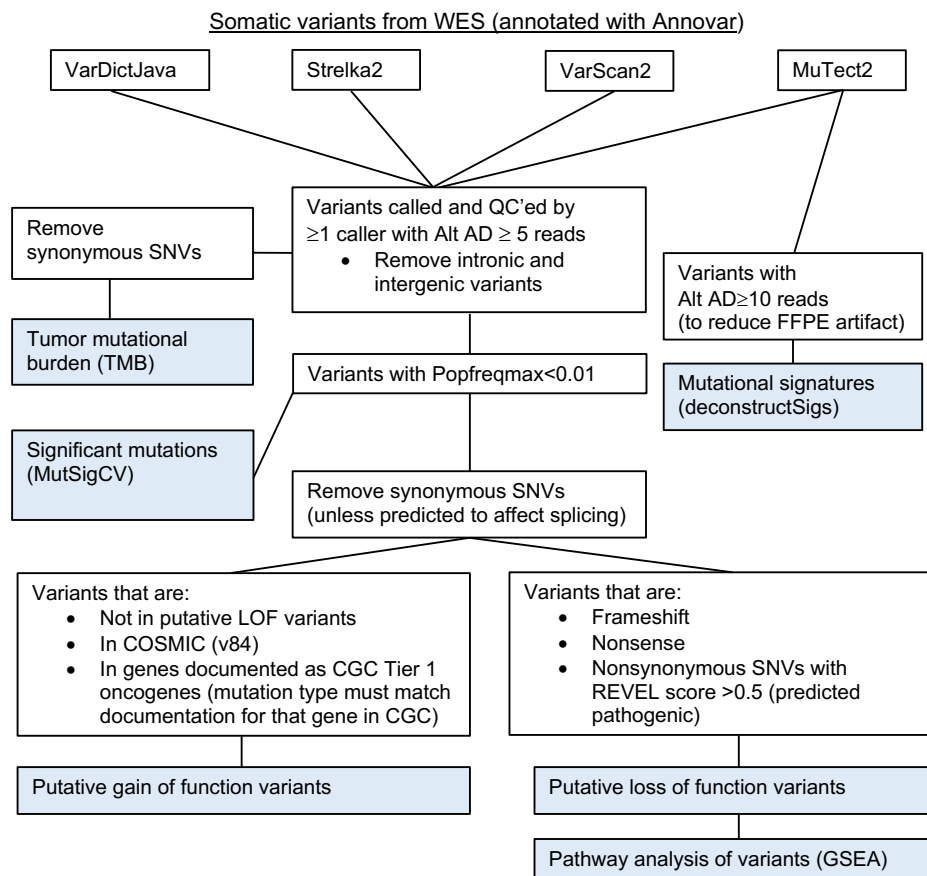


Figure 5: Manual filtering strategy for somatic variants. Unfiltered somatic variants from four variant callers were manually filtered for various downstream analyses as indicated. Alt AD refers to alternative allele depth.

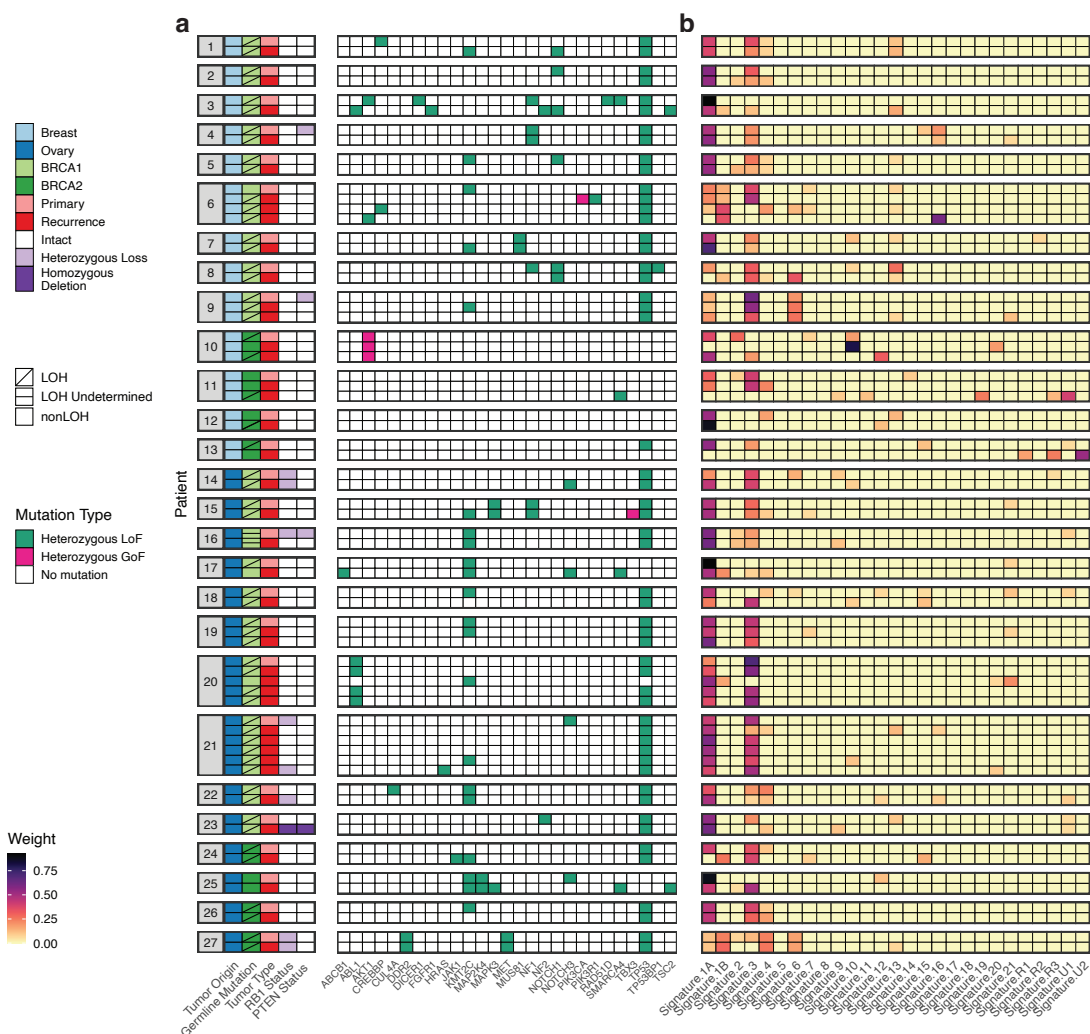


Figure 7: Somatic mutations and mutational signatures by tumor. A. Phenotype, *BRCA1/2* allele-specific LOH, *RB1* and *PTEN* status, and somatic mutations found in 67 paired primary and recurrent tumors from 27 patients sequenced by WES (n=67) and high-depth targeted sequencing (n=44). Mutations from WES and targeted sequencing are merged. Tumors for each patient are displayed in chronological order (top to bottom). Display is limited to genes with ≥ 1 mutation with alternative allele fraction ≥ 0.05 from targeted sequencing. B. Mutational signatures derived from WES mutations.

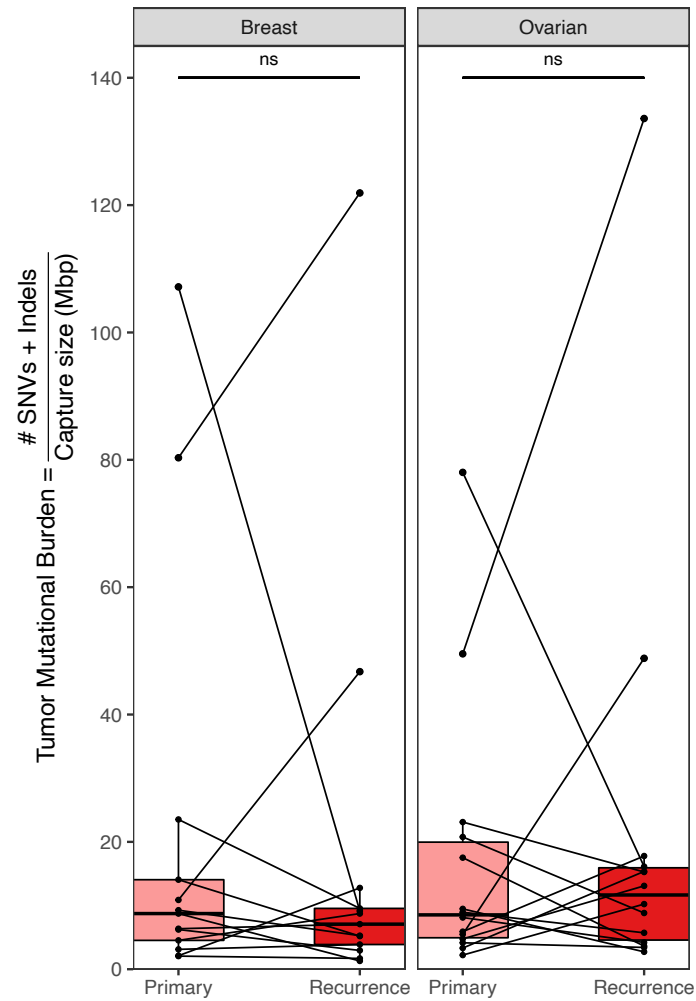


Figure 8: Paired comparisons of tumor mutational burden. Comparison of tumor mutational burden (calculated from WES mutations with alternative allele depth ≥ 5 reads) for primary/recurrent tumor pairs. For patients with multiple recurrences, one recurrence was chosen at random for comparison. Pairwise differences were determined by two-sided Wilcoxon signed rank test ($\alpha=0.05$).

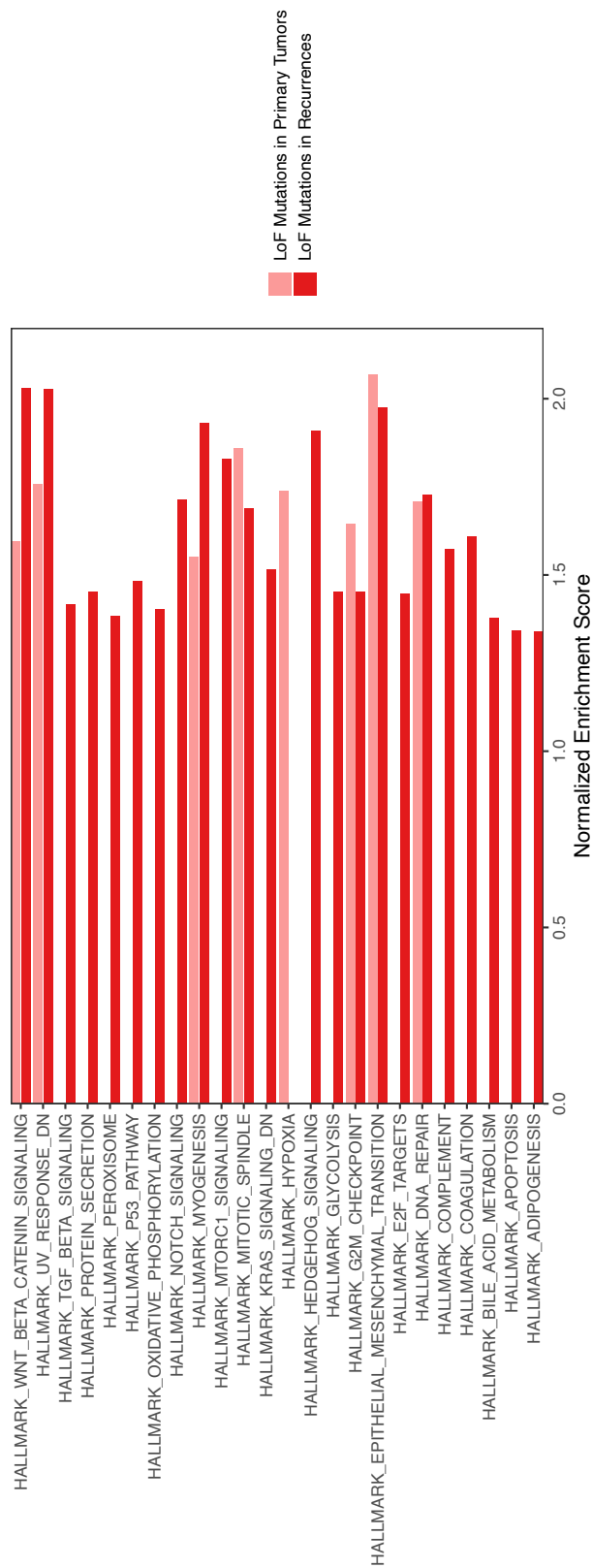


Figure 9: Pathway analysis of loss of function mutations. Hallmark Gene Sets enriched in loss of function mutations from WES, computed separately for primary and recurrent tumor groups (all FDR $q < 0.25$).

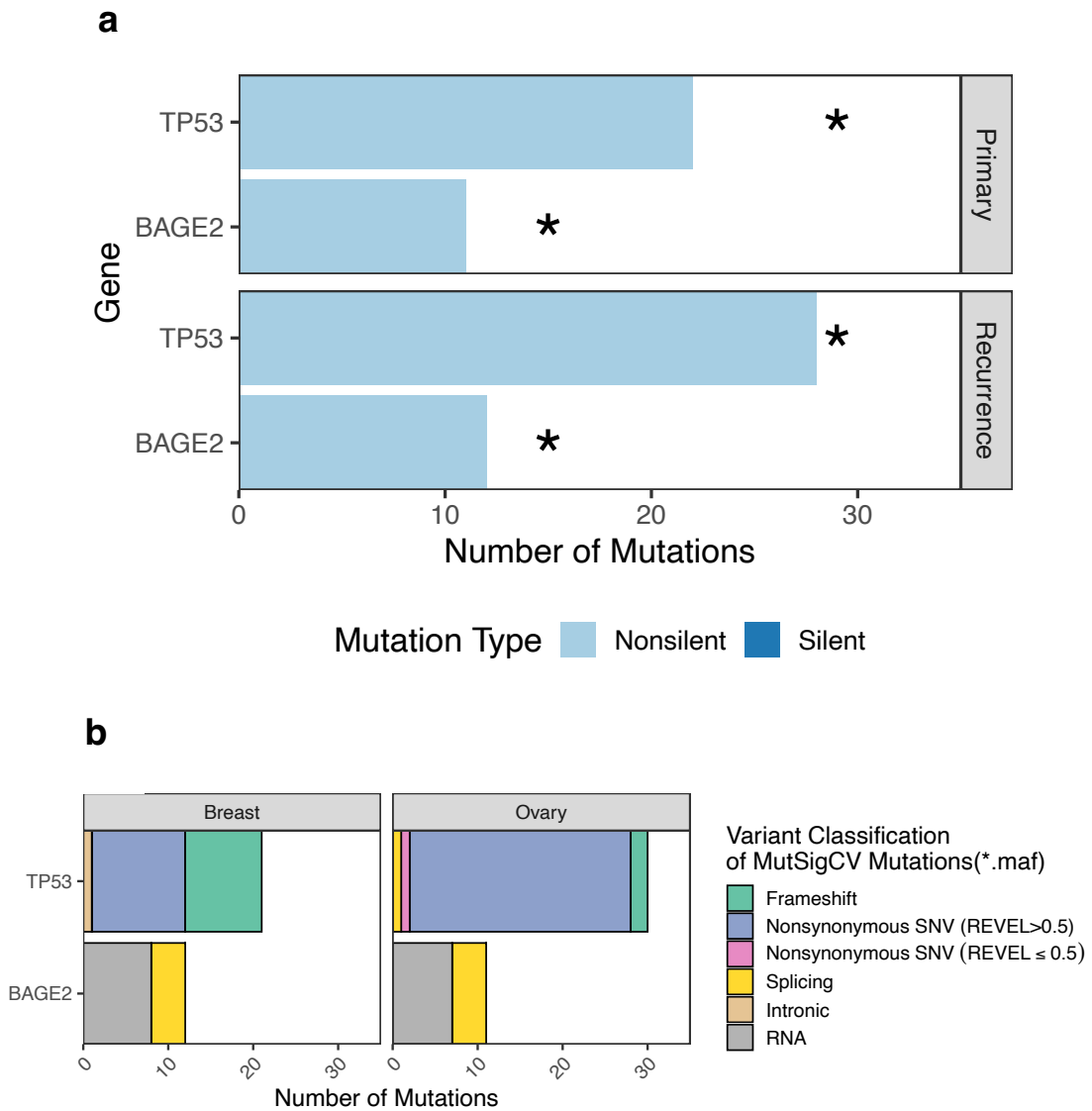


Figure 10: Significant mutations in individual genes. A. MutSigCV results from WES (*FDR $q < 0.05$). B. Variant classification of mutations contributing to MutSigCV results in A, by tumor type. Variant classifications correspond to the exact variants (*.maf) used for MutSigCV analysis.

Discussion

Using a relatively large sample set of paired primary and recurrent tumors, we interrogated the evolution of somatic mutations in *BRCA1/2* mutation-associated breast and ovarian cancer. Our study represents the largest comprehensive assessment of somatic mutations in these tumors to date. We found that most malignancies were driven by LoF mutations in *TP53*, in agreement with previous findings^{31,46,47,100–102}. However, some tumors (particularly ER+ breast tumors in *BRCA2* mutation carriers) did not have *TP53* mutations and were likely driven by alternative mechanisms (including but not limited to the oncogenic *AKT1* E17K mutation). In addition to *TP53*, our MutSigCV analysis also yielded the lncRNA *BAGE2* as a significantly mutated gene. While *BAGE2* is not necessarily a tumor driver, this result suggests that it could be a novel tumor antigen not previously observed in breast or ovarian cancer.

Surprisingly, we found that tumor mutational burden was not significantly higher in recurrent tumors compared to matched primary tumors. This finding conflicts with an earlier study positing that ovarian tumor TMB increases after platinum treatment, but our study included 14 primary/recurrent ovarian tumor pairs in comparison to their smaller sample size of five tumor pairs⁵¹. To our knowledge, this is the first assessment of TMB in paired primary/recurrent *BRCA1/2* mutation-associated breast tumors. Overall, TMB results suggest that *BRCA1/2* mutation-associated tumors accumulate the majority of their mutations early in the process of tumorigenesis, and that mounting mutations are unlikely to represent the main mechanism for recurrence. This result is also supported by prior studies, which found high concordance between somatic mutations in

unselected primary and recurrent breast and ovarian tumors (87% concordance and >51% concordance, respectively)^{103–105}.

Our mutational signatures analysis further suggests that most somatic mutations result from HR deficiency (due to partial or complete *BRCA1/2* loss), as well as aging. Although *BRCA1/2* mutation carriers do tend to develop their cancers earlier than sporadic cases, most are not diagnosed until middle age¹⁰. Depending on the cohort, it is estimated that the mean age of breast or ovarian cancer diagnoses in *BRCA1/2* mutation carriers can range from 42 to 61 years of age^{12,14,106}. Of note, breast cancer is diagnosed earlier than ovarian cancer in *BRCA1/2* mutation carriers (forties vs. fifties/sixties), but aging-associated signatures drove somatic variants in both tumor types^{12,14,106}.

We assessed the pathways most commonly affected by LoF mutations, as a potential readout of dysregulated signaling in primary and recurrent tumors. Mutations in recurrences affected a larger number of gene sets (23) than mutations in primary tumors (8). Most (7/8) pathways that were mutated in primary tumors were also mutated in recurrences, suggesting that much of pro-tumorigenic signaling facilitated by mutations in a primary tumor is conserved upon recurrence. Gene sets that were exclusively mutated in recurrences included additional pathways for growth, survival, and metabolic processes.

We interrogated the mutation status of five genes associated with platinum and PARPi resistance in the literature: *ABCB1*, *TP53BP1*, *MAD2L2*, *CHD4*, and *PARP1*^{23,51,66–68,70,71,73,99}. Activity of the ABC transporter MDR1 has been shown to mediate multidrug resistance^{51,71}. MDR1 activity is commonly facilitated by gene fusions

that connect the encoding gene (*ABCB1*) to an active promoter to increase *ABCB1* mRNA expression^{51,71}. Accordingly, we did not expect mutations to represent a major mechanism to *ABCB1* hyperactivation. While we did find one potential GoF mutation in *ABCB1*, the presence of a nearby pathogenic mutation in the same gene may have abrogated any resulting MDR1 activation. We also did not identify pathogenic mutations in *CHD4* or *MAD2L2/REV7*, suggesting that mutations in *CHD4*, *MAD2L2/REV7*, and *ABCB1* are not a major mechanism of acquired therapeutic resistance in this cohort.

We identified one pathogenic *TP53BP1* mutation in the entire cohort, in a breast cancer recurrence from patient 8. Patient 8 was a *BRCA1* mutation carrier who received extensive treatment prior to recurrence, including platinum, PD-1/PARPi combination therapy, hormonal therapy, chemotherapy, and radiation (see **Supplementary File 1** for complete treatment history). Her primary tumor was treatment-naïve and lacked the pathogenic *TP53BP1* mutation. *TP53BP1* loss has been reported previously as a PARPi resistance mechanism, by which *BRCA1*-deficient cells reprogram the DNA damage response to perform *BRCA1*-independent HR^{67,68,99}. While this effect has been demonstrated *in vitro* and *in vivo*, this is the first published report of somatic *TP53BP1* loss in a *BRCA1* mutation carrier after platinum and PARPi.

We identified one pathogenic *PARP1* mutation in the entire cohort, in an ovarian cancer recurrence from patient 20. Patient 20 was a *BRCA1* mutation carrier who received platinum/taxane chemotherapy, as well as doxil, prior to this recurrence. Since *PARP1* loss has only been reported to mediate PARPi resistance *in vitro*, it is unclear whether the tumor derived any benefit from this *PARP1* mutation⁷³.

Lastly, one major goal of our somatic mutation analysis (and much of the

rationale for performing targeting sequencing) was to identify *BRCA1/2* reversion mutations in recurrences after treatment. We were surprised to discover that only two out of 40 recurrences showed evidence of *BRCA1/2* reversion. These two recurrences were from the same *BRCA1* mutation carrier, Patient 20, and had the same in-frame deletion (see chapter 7). Conversely, no reversions were identified for any *BRCA2* mutation carriers, or any *BRCA1* mutation carriers with breast cancer. Ultimately, 38 of 40 recurrences (95%), including 17 post-platinum recurrences and four post-platinum+PARPi recurrences, did not have *BRCA1/2* reversion mutations. Our findings agree with a recent meta-analysis of *BRCA1/2* reversion mutation studies, which found that *BRCA1/2* reversions vary in prevalence based on germline mutation type and position⁶⁰. This study also found that most reversion mutations are small somatic deletions around the site of a germline truncating mutation, as was the case in Patient 20⁶⁰. Overall, our results support the conclusion that *BRCA1/2* reversion mutations do not represent a main (or even common) mechanism of acquired therapeutic resistance to platinum, PARPi, or other conventional therapies in *BRCA1/2* mutation-associated cancers. We also found that somatic *BRCA1/2* mutations do not constitute a mechanism of *BRCA1/2* loss of heterozygosity in this cohort either.

CHAPTER 4: ANALYSIS OF COPY NUMBER VARIATION

Introduction

BRCA1/2 mutation-associated tumors typically demonstrate somatic loss of the wild-type *BRCA1/2* allele, a process referred to as “loss of heterozygosity,” or LOH. Tumors with *BRCA1/2* LOH are mostly or completely deficient for *BRCA1/2*, having lost one allele to the pathogenic germline mutation and the other allele to somatic events. Tumors deficient for *BRCA1/2* are also correspondingly deficient for homologous recombination and sensitive to platinum and PARPi^{37–44}. Given that LOH has been observed in normal breast tissue from *BRCA1/2* mutation carriers, it is thought to be an early (but not required) event in tumorigenesis^{32,33}. LOH is most commonly mediated by copy number losses of the wild-type *BRCA1/2* allele, but can also arise from deleterious somatic mutations or other mechanisms³¹. However, as mentioned in chapter 3, we did not observe somatic mutations causing *BRCA1/2* LOH in this cohort.

For many years, it was assumed that all *BRCA1/2* mutation-associated tumors underwent LOH^{35,36}. Conversely, our lab showed that a minority of tumors do not undergo LOH, instead retaining the wild-type *BRCA1/2* allele³¹. Tumors without LOH (termed “nonLOH” here) were most common in *BRCA2* mutation carriers with breast cancer (46% of tumors)³¹. NonLOH tumors were present, but less common, in cohorts of *BRCA1* mutation-associated breast cancers (10% of tumors) and *BRCA1/2* mutation-associated ovarian cancers (7% and 16% of tumors, respectively)³¹. Since *BRCA1/2* proficiency has implications for therapeutic response and survival, we used allele-specific copy number to investigate *BRCA1/2* LOH status in 67 paired primary and recurrent *BRCA1/2* mutation-associated tumors³¹. We aimed to assess whether LOH status changed or remained consistent over the course of disease, especially after

treatments that exploit deficiencies in BRCA1/2 and HR (platinums, PARPi).

In parallel with our LOH analyses, we used several copy number-based metrics to quantify genomic abnormalities in these tumors. Composite HRD scores have been used previously to enumerate the chromosomal scarring specific to HR-deficient backgrounds^{31,107–109}. These scores consist of three metrics: large-scale state transitions, non-telomeric allelic imbalance, and loss of heterozygosity^{110–112}. (Here, loss of heterozygosity does not refer to *BRCA1/2* status, but rather to locus-agnostic segments of LOH across the genome¹¹⁰.) Each metric represents an orthogonal approach to quantifying HRD: large-scale state transitions and genomic LOH reflect *BRCA1/2* germline mutation status, while non-telomeric allelic imbalance predicts platinum sensitivity^{110–112}. An unweighted sum of these three metrics was shown to predict platinum sensitivity and *BRCA1/2* germline mutation status in breast cancer, and we employed this approach for our study (see **Methods**)¹⁰⁷. Of note, HRD scores also differ based on *BRCA1/2* LOH status, with nonLOH tumors scoring lower than tumors of the same type with LOH³¹. We also interrogated aneuploidy in this sample set, as a second approach to quantifying chromosomal aberrations.

Lastly, we sought to determine the functional effects of copy number variation on specific genes and pathways in paired primary and recurrent tumors. We employed two orthogonal methods of gene annotation in segments of copy number variation (CNVs). Briefly, we performed segmentation analysis with Sequenza, which defines CNVs based on tumor vs. germline DNA read depth comparisons across the genome¹¹³. First, Sequenza-derived segments for each tumor were binned based on total copy number and annotated with genes (see **Methods**). Using this threshold-based approach, we assessed the effects of CNVs on pathways (by tumor group) and on individual genes of

interest (by tumor). Secondly, we performed GISTIC analysis on Sequenza-derived segments, to identify genes that were significantly amplified or deleted by tumor group¹¹⁴. Using one method that prioritizes total copy number regardless of segment length (binning and annotation) and another method that prioritizes segment length regardless of total copy number (GISTIC), we hoped to improve the confidence of our results^{113,114}.

Methods

Genome-wide analysis of copy number variation

We performed allele specific copy number analysis using Sequenza (v3.0), to identify segments of copy number variation (CNVs) as well as ploidy and cellularity estimates¹¹³. Segments were used to calculate homologous recombination deficiency (HRD) and aneuploidy scores, using custom scripts and as previously described (<https://github.com/maxwell-lab/HRDex>)^{31,115}. We used GISTIC2.0 (v2.0.23, Gene GISTIC mode, with default cutoffs) to identify significant CNVs at 90% confidence in tumor cohorts¹¹⁴. Genes were highlighted from segments with residual $q \leq 0.05$. BEDtools (v2.28.0) was used to compute Jaccard statistics for similarity between CNV sets, limited to segments of $\geq 50\%$ reciprocal overlap¹¹⁶. CNV segments were binned based on total (integer) copy number using the following convention: CN=0, Deletion; CN=1, Loss; CN=2-3, Neutral; CN=4-5, Gain; CN>5, Amplification.

Gene-level annotation of copy number variation

Binned segments were annotated with genes using AnnotSV (v2.3)¹¹⁷. Genes were considered to be lost or deleted if >50% of the locus was encompassed by a copy number loss or deletion. Conversely, genes were only considered to be gained or

amplified if 100% of the locus was encompassed by a gain or amplification. In the case of a gene spanning several different segments of CNV, the copy number of the gene was reported as the minimal copy number of the entire locus.

Pathway analysis of genes subject to copy number variation

For pathway analysis of genes subject to copy number gains, genes were ranked based on frequency of their presence in a segment of gain or amplification, within primary and recurrent tumor groups. Similar to our GSEA Preranked analysis of variants (see Chapter 3), a gene could only be counted as gained once per patient per tumor group. We then repeated this process to rank genes again based on frequency of copy number loss or deletion by tumor group. Ranked gene lists of “gained/amplified” genes and “lost/deleted” genes were used as input for GSEA Preranked with Hallmark gene sets (1000 permutations)⁹³. Gene set enrichment was considered significant using a cutoff of $FDR < 0.25$.

Detection of *BRCA1/2* biallelic loss by copy number variation

We assessed loss of the wild-type *BRCA1* or *BRCA2* allele via allele-specific copy number variation from Sequenza. Briefly, we used tumor and germline bam files from WES for Sequenza input. Zygosity was estimated for all segments of somatic copy number variation based on the number of A (major) and B (minor) alleles detected. Segments with zero B alleles (only A alleles present) were deemed to have undergone LOH. Conversely, if at least one B allele was detected, a segment was considered nonLOH (heterozygous). Tumors were deemed to have undergone loss of heterozygosity (LOH) if the segment containing their pathogenic *BRCA1/2* germline mutation had LOH. When possible, we confirmed proficient *BRCA1/2* expression in

nonLOH tumors with RNA sequencing. For patients with LOH transitions, we assessed the allele fraction of all *BRCA1/2* germline variants (SNPs and the pathogenic variant) across tumors to further corroborate LOH calls.

Acquisition of validation cohorts from TCGA PanCancer datasets

For validation of *PARP1* copy number gains, we used WES from primary breast and ovarian tumors in TCGA cohorts (**Supplementary File 7**). Level 1 WES bam files from TCGA Breast and Ovarian cohorts were obtained from Genomic Data Commons (<https://gdc.cancer.gov/>) as previously described^{31,115}. Tumors were excluded from analysis for any of the following reasons: (1) failed quality control assessment; (2) had somatic *BRCA1/2* loss as indicated by *BRCA1/2* mutation, *BRCA1/2* deletion, low *BRCA1/2* expression, or *BRCA1* promoter methylation; or (3) had germline or somatic mutations in genes related to homologous recombination (see **Supplementary File 7** for HR gene list). TCGA tumors were assessed for these features, and grouped in g*BRCA1/2* or HR-WT groups, as previously described and visualized in **Figure 11**^{31,115}.

Statistical analysis

All paired comparisons of HRD and aneuploidy scores were tested for significance with two-sided Wilcoxon signed rank tests ($\alpha=0.05$). For patients with multiple recurrences, one recurrence was chosen at random for pairwise comparisons (such that $n=27$ for each group). In the primary/recurrent cohort, groupwise differences in average *PARP1* copy number were assessed by Kruskal-Wallis test, with significant results followed by Dunn's test with Bonferroni correction ($\alpha=0.05$). In TCGA cohorts, groupwise differences in average *PARP1* copy number were assessed by two-sided t-tests ($\alpha=0.05$). *PARP1*

copy number is reported for primary/recurrent and TCGA tumors in **Supplementary File 6**.

Results

Analysis of *BRCA1/2* loss of heterozygosity by copy number variation

Using (WES-derived) allele-specific copy number from Sequenza, we assessed *BRCA1/2* allele-specific loss of heterozygosity (LOH) across the cohort of primary and recurrent tumors. Since we did not find secondary somatic mutations facilitating this loss (see Chapter 3), we interrogated whether tumors underwent LOH via copy number loss of their wild-type *BRCA1/2* allele (see **Methods**)¹¹³. Overall, 22/27 (81%) primary tumors and 33/40 (82%) recurrences exhibited this mechanism of LOH (**Fig. 12**). Of the 10 “nonLOH” tumors (with retention of the wild-type *BRCA1/2* allele), five were *BRCA2* mutation-associated breast tumors. These results are consistent with prior findings that absence of LOH is more common in *BRCA2* mutation-associated tumors³¹. Notably, we found that most primary/recurrent tumor pairs had consistent *BRCA1/2* allele-specific LOH status.

We observed discordant LOH status more frequently in breast cancers (five – 38%) than in ovarian cancer patients (two – 14%). In particular, three out of four ER+ breast cancer patients in the cohort had discordant LOH in their tumors. LOH transitions were also observed for nearly every patient with *TP53* wild-type tumors (6/7 patients – 86%). In three patients (Patients 6, 11, and 12), the primary tumor did not have LOH, but first exhibited LOH upon recurrence (**Fig. 13**). Conversely, Patients 10, 13, 17, and 20 showed the opposite effect: their primary tumors exhibited LOH but developed into a nonLOH recurrence (**Fig. 14**).

This “LOH reversal” was observed in two *BRCA2* mutation carriers with breast cancer (treated with chemotherapy and radiation) and two *BRCA1* mutation carriers with ovarian cancer (treated with platinum). Mutational signatures reflected the change in HR proficiency (and potential for other drivers) among tumors with LOH reversal; loss of Signature 3 was observed in patients 10, 13, and 20 (**Fig. 7b**). In Patients 10, 17, and 20, nonLOH recurrences also had a lower HRD score compared to matched primary tumors with LOH (**Fig. 15**). The fourth patient with LOH reversal (Patient 13) showed copy number losses and a LoF mutation in key DNA repair genes upon recurrence, which could have contributed to a high HRD score even in the context of *BRCA2* proficiency. In contrast, for all three tumors that developed LOH upon recurrence, nonLOH primary tumors had a lower HRD score than recurrences (**Fig. 15**). Comparisons of HRD scores between tumors with discordant LOH status were not significant, likely due to small sample size (3-4 tumors). Overall, we found that LOH discordance was more common in patients with ER+ breast cancer and those without somatic *TP53* LoF mutations. We did not find any other associations between LOH discordance and patient clinical correlates in this limited sample set.

Assessment of genomic scarring and copy number variation

We next evaluated whether recurrent tumors showed more chromosomal instability and genomic scarring than primary tumors. Neither aneuploidy nor HRD scores (individual metrics or combined), were significantly different between tumor pairs overall, although some primary and recurrent tumors had quite different scores (**Fig. 16; Supplementary File 2**). To identify commonly amplified and deleted genes in primary and recurrent tumors, we performed GISTIC analysis, stratified by tumor type (**Supplementary File 5**)¹¹⁴. We assessed individual genes that fell within significant

amplifications or deletions (90% CI and FDR $q < 0.05$) and were biologically relevant to breast and ovarian cancer. Both tumor types demonstrated significant *CHD2* and *MYC* amplifications (**Fig. 17**). In a comparison of GISTIC segments between breast and ovarian tumors, we identified intersections (>50% overlap) for 16 deletions and eight amplifications. Breast tumors demonstrated amplifications of *SOX4*, which has been shown to promote PI3K/AKT signaling (**Fig. 17a**)¹¹⁸. We also observed deletions of putative breast tumor suppressor *RUNX3*, DNA repair genes *RAD18* and *XPC*, and chromatin remodelers associated with cancer (*HDAC4/10/11* and *BRD1*) (**Fig. 18a**)^{119–121}. Specific to ovarian tumors were amplifications in the ABC transporter *ABCB11*, for which gains are associated with chemoresistance *in vitro* (**Supplementary Fig. 17b**)¹²². Other notable ovarian-specific deletions included DNA damage response genes (*ERCC1/2*, *CHEK1*, *FANCF*), apoptotic regulators (*BAX*, *BAD*), *PTEN*, *WEE1*, and *CDKN1C* (**Fig. 18b**).

As the tumor types exhibited similar copy number changes, we combined all primary and all recurrent tumors separately to identify commonalities. Primary and recurrent tumor groups each were significantly enriched for distinct regions of amplifications and deletions. Significant (90% CI and FDR $q < 0.05$) deletions were more similar between groups than significant amplifications (Jaccard similarity index of 0.06 vs. 0.20 for segments of 50% overlap). Primary tumors had amplifications in *MYC* and *CHD2* (**Fig. 19a**). Recurrent tumors shared these amplifications and were also uniquely enriched for amplifications in *PARP1* (**Fig. 19b**). Recurrences also showed amplifications of *ABCB10*, an ABC transporter for which copy number gains are associated with chemotherapy resistance *in vitro*¹²². Further, we noted recurrence-specific amplifications in *TGFB2*, *TRAF5*, *WNT3A*, *SOX4*, and *E2F3*, suggesting that

recurrences may benefit from increased dosage of genes that regulate growth and lineage signals. Unique, significant deletions suggested that primary tumors undergo loss of tumor suppressor genes (*ESR1*, *VHL*) and changes in chromatin accessibility (*HDAC10*, *HDAC11*, *BRD1*) (**Fig. 20a**). In recurrences, unique deletions included *CASP9*, *CDKN2C*, *JUN*, *MUTYH*, and *RAD54L*, suggesting dysregulation of cell cycle, DNA damage response, and apoptosis (**Fig. 20b**).

Copy number analysis of *PARP1* in primary/recurrent and TCGA cohorts

We found *PARP1* to be significantly (90% CI, FDR $q=0.05$) amplified in recurrences but not primary tumors. Given the use of PARPi for treatment of late-stage *BRCA1/2* mutation-associated tumors, we investigated further. Over half of tumors (35/67, 52%) in the cohort had copy number gains (CN=4-5) or amplifications (CN>5) in *PARP1* (**Supplementary File 6**). Although *PARP1* gains and amplifications were present in primary breast (62%), primary ovarian (43%), and recurrent ovarian tumors (32%), the finding was most frequent in breast recurrences (78%) (**Fig. 21**). Seventy percent of all patients (19/27) in the cohort had a *PARP1* gain or amplification in at least one tumor (85% of breast cancer patients, 57% of ovarian cancer patients). Most patients (11/19, 60%) had a *PARP1* gain or amplification in both primary and recurrent tumors, 5/19 patients gained *PARP1* upon recurrence, and 3/19 patients had a *PARP1* gain or amplification exclusive to their primary tumor. Of six recurrences evaluated after treatment with PARPi, three had a *PARP1* gain, whereas three were copy neutral for *PARP1*. Gains and amplifications tended to be focal and centered around the *PARP1* locus, rather than arm-level (**Fig. 22**), underscoring a potential selection effect¹¹⁴. Conversely, we noticed *PARP1* copy number losses in only 2/27 (7%) patients. In particular, three ovarian recurrences from a single *BRCA1* mutation carrier underwent

PARP1 loss (1 copy) and one breast recurrence from a *BRCA2* mutation carrier underwent *PARP1* deletion (0 copies).

We next evaluated *PARP1* copy number variation in a larger, independent sample set to validate our findings. Specifically, we assessed whether *PARP1* gains and amplifications were more prevalent in *BRCA1/2* mutation-associated tumors compared to sporadic breast and ovarian tumors from TCGA PanCancer cohorts. We grouped TCGA (primary) tumors into *BRCA1/2* germline mutation-associated (*gBRCA1/2*) and HR wild-type (HR-WT) groups, as previously described (**Fig. 11, Supplementary File 7**)^{31,115}. We assessed *PARP1* copy number in these cohorts, replicating the Sequenza analysis performed in the primary/recurrent cohort. We found a high prevalence of *PARP1* gains and amplifications in *gBRCA1/2* and sporadic breast and ovarian tumors, with no differences between *gBRCA1/2* and sporadic tumors of the same type (**Fig. 23, Supplementary File 6**).

Pathway analysis of genes affected by copy number variation

Finally, we performed GSEA Preranked to assess which pathways were affected by Sequenza-derived copy number gains and losses (**Fig. 24**). As in the GSEA analysis of LoF mutations, we performed GSEA Preranked on lists of genes ranked by prevalence of copy number gain or loss within each tumor group (**Supplementary File 3**). Primary tumors were enriched for gains in gene sets for UV response, $\text{TNF}\alpha$ signaling via $\text{NF-}\kappa\text{B}$, $\text{IFN}\alpha$ and $\text{IFN}\gamma$ response, and angiogenesis. Recurrent tumors were uniquely enriched for gains in *MTORC1* signaling genes and deletions in Hedgehog signaling genes, consistent with our finding that recurrences accumulate LoF mutations in the Hedgehog signaling gene set.

Copy number variation in genes related to platinum and PARPi resistance

Lastly, we interrogated the effects of WES-derived copy number variation on genes associated with platinum and PARPi resistance in the literature (outside of *PARP1*, discussed above). First, we looked for copy number gains and amplifications in *ABCB1*, which could facilitate increased expression of the encoded efflux pump. We found copy number gains in *ABCB1* (ranging from 4-8 copies) in 12 primary tumors and 20 recurrences. Eleven of 12 (92%) patients with an *ABCB1* gain in their primary tumor retained the gain upon recurrence, while four patients had recurrence-exclusive *ABCB1* gains.

Next, we looked for copy number losses or deletions in *REV7/MAD2L2*, *TP53BP1*, and *CHD4*. We specifically looked for *CHD4* loss in *BRCA2* mutation-associated tumors, and *REV7/MAD2L2* and *TP53BP1* loss in *BRCA1* mutation-associated tumors, to reflect the genetic backgrounds in which these losses mediated platinum and/or PARPi resistance *in vitro*^{66–68,70,99}. *CHD4* was deleted in a single *BRCA2* mutation carrier's recurrence following chemotherapy. *REV7/MAD2L2* was lost or deleted in three *BRCA1* mutation-associated breast tumors (two primary tumors and one matching recurrence), as well as five *BRCA1* mutation-associated ovarian recurrences. Of note, one of these ovarian recurrences was platinum- and PARPi-treated (Patient 19, recurrence #2). *TP53BP1* loss was detected in eight tumors (four primary and four recurrent) from two *BRCA1* mutation carriers with breast cancer and four with ovarian cancer. Of note, all four *BRCA1* mutation-associated recurrences with *TP53BP1* losses were pre-treated with platinum. One of these was the same post-platinum, post-PARPi ovarian tumor observed to have *REV7/MAD2L2* loss (Patient 19, recurrence #2).

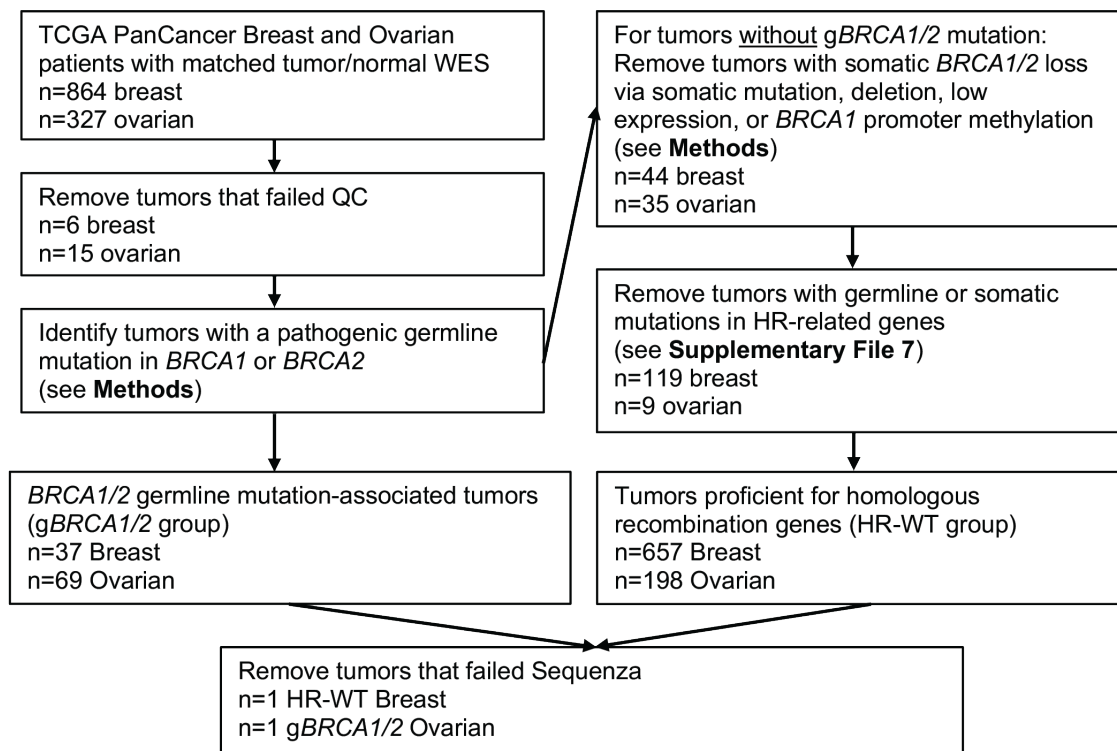


Figure 11: Filtering strategy for TCGA PanCancer cohorts. Strategy for grouping TCGA breast and ovarian tumors into *BRCA1/2* germline mutation-associated (“g*BRCA1/2*”) and homologous recombination wild-type (“HR-WT”) groups for analysis. Tumors were excluded on the basis of somatic *BRCA1/2* loss (by any of multiple mechanisms) or germline or somatic mutations in genes required for HR. Tumors and HR genes are listed in **Supplementary File 7**.

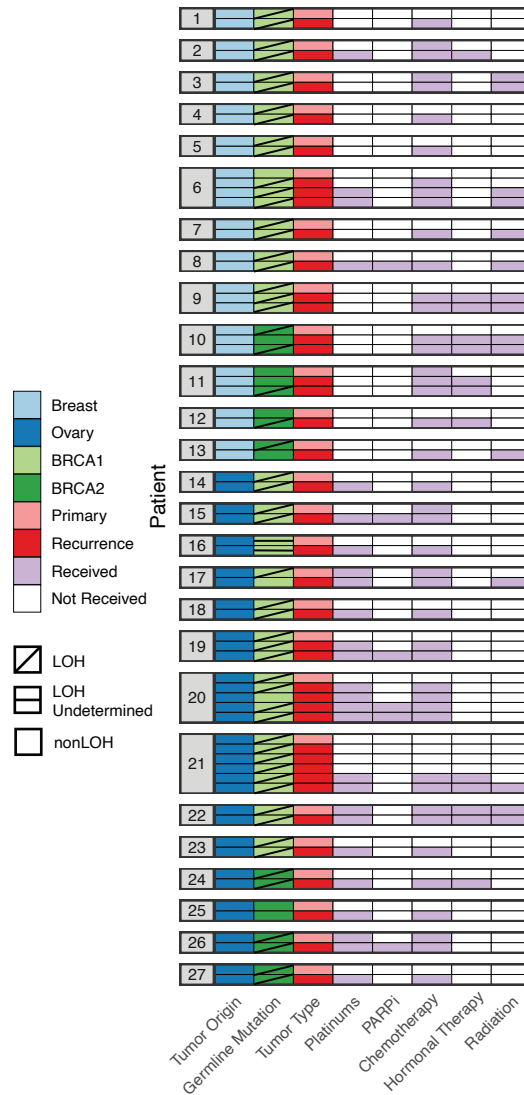


Figure 12: *BRCA1/2* loss of heterozygosity by tumor. Allele-specific loss of *BRCA1/2* heterozygosity (lines) in 67 paired primary and recurrent tumors. Tumors were exposed to treatment prior to resection as indicated. Tumors for each patient are displayed in chronological order (top to bottom).

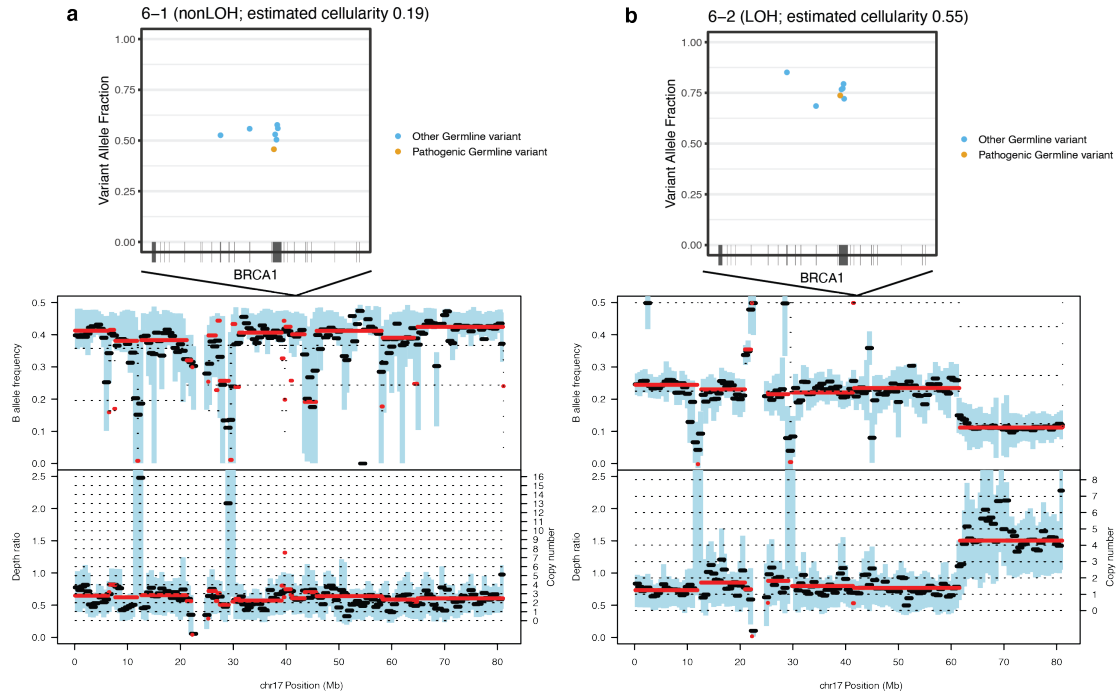


Figure 13: *BRCA1* germline variants, B allele frequency, and copy number before and after a representative nonLOH to LOH transition. A. Location and variant allele fraction of *BRCA1* germline variants in Patient 6's primary breast tumor without LOH (top), pictured with B allele frequency and copy number across chromosome 17 (bottom). B. Location and variant allele fraction of *BRCA1* germline variant in Patient 6's first breast tumor recurrence with LOH (top), pictured with B allele frequency and copy number across chromosome 17 (bottom). For A and B, germline variants were called by VarDictJava and VarScan2. B allele frequency and copy number plots were generated in Sequenza.

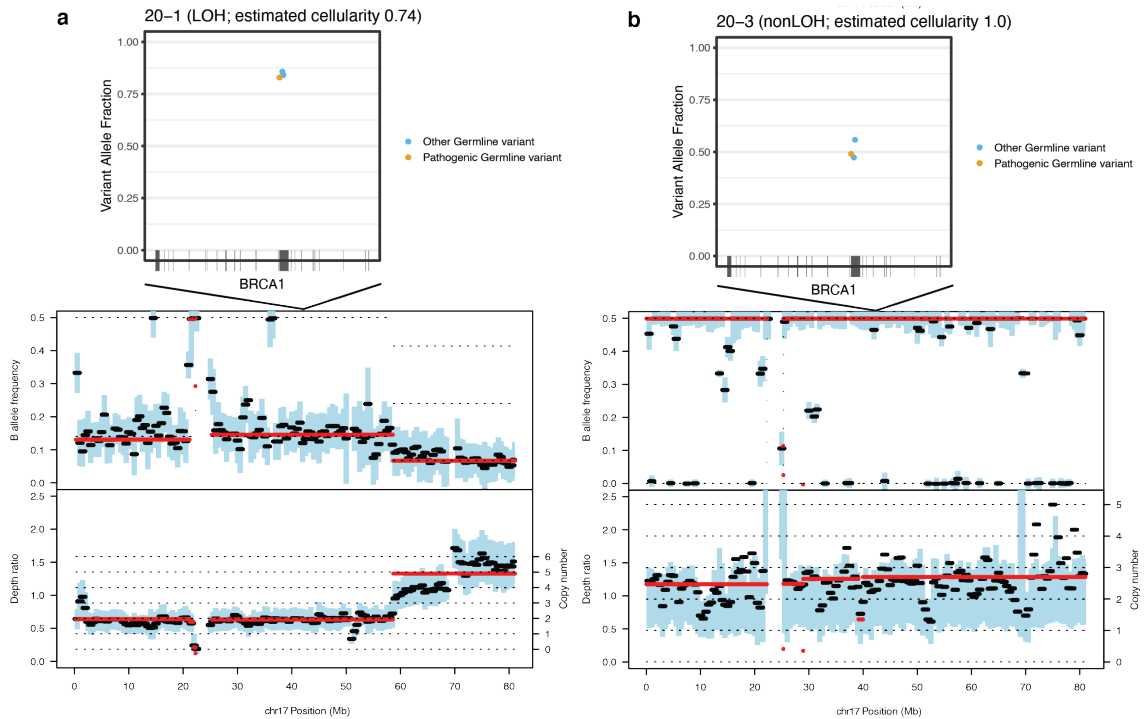


Figure 14: *BRCA1* germline variants, B allele frequency, and copy number before and after a representative LOH reversal. A. Location and variant allele fraction of *BRCA1* germline variants in Patient 20's primary ovarian tumor with LOH (top), pictured with B allele frequency and copy number across chromosome 17 (bottom). B. Location and variant allele fraction of *BRCA1* germline variant in Patient 20's second ovarian tumor recurrence without LOH (top), pictured with B allele frequency and copy number across chromosome 17 (bottom). For A and B, germline variants were called by VarDictJava and VarScan2. B allele frequency and copy number plots were generated in Sequenza.

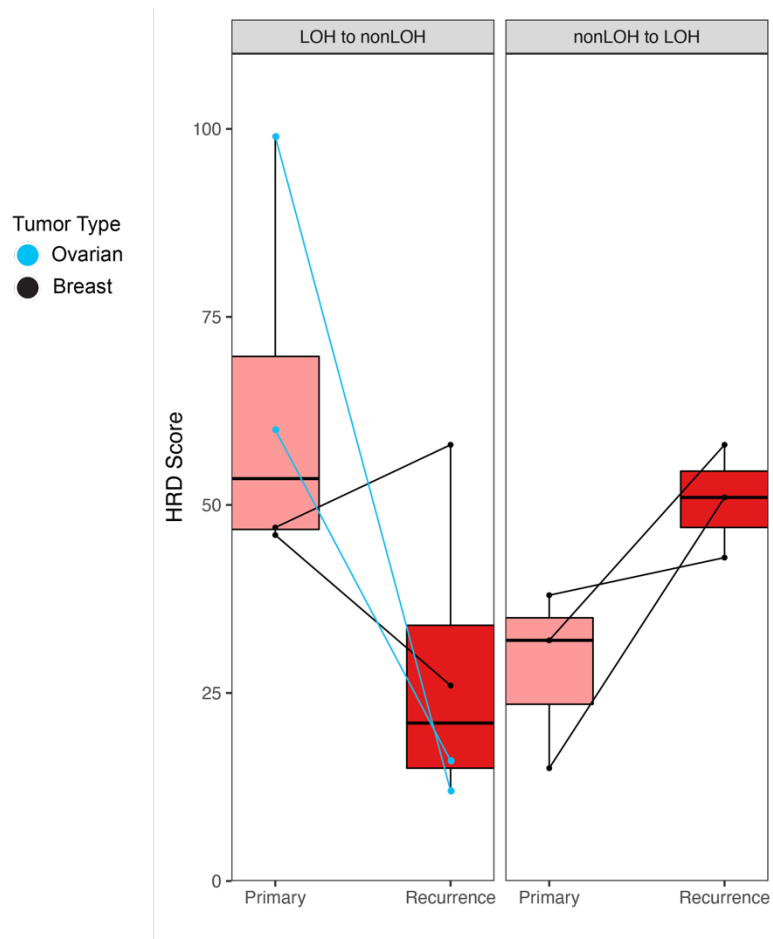


Figure 15: HRD score comparisons in tumors with *BRCA1/2* LOH transitions.

Comparison of HRD score for tumors with a change in LOH status from primary to recurrence. For one patient (Patient 6), who had multiple recurrences with LOH, the first recurrence after nonLOH→LOH transition is displayed.

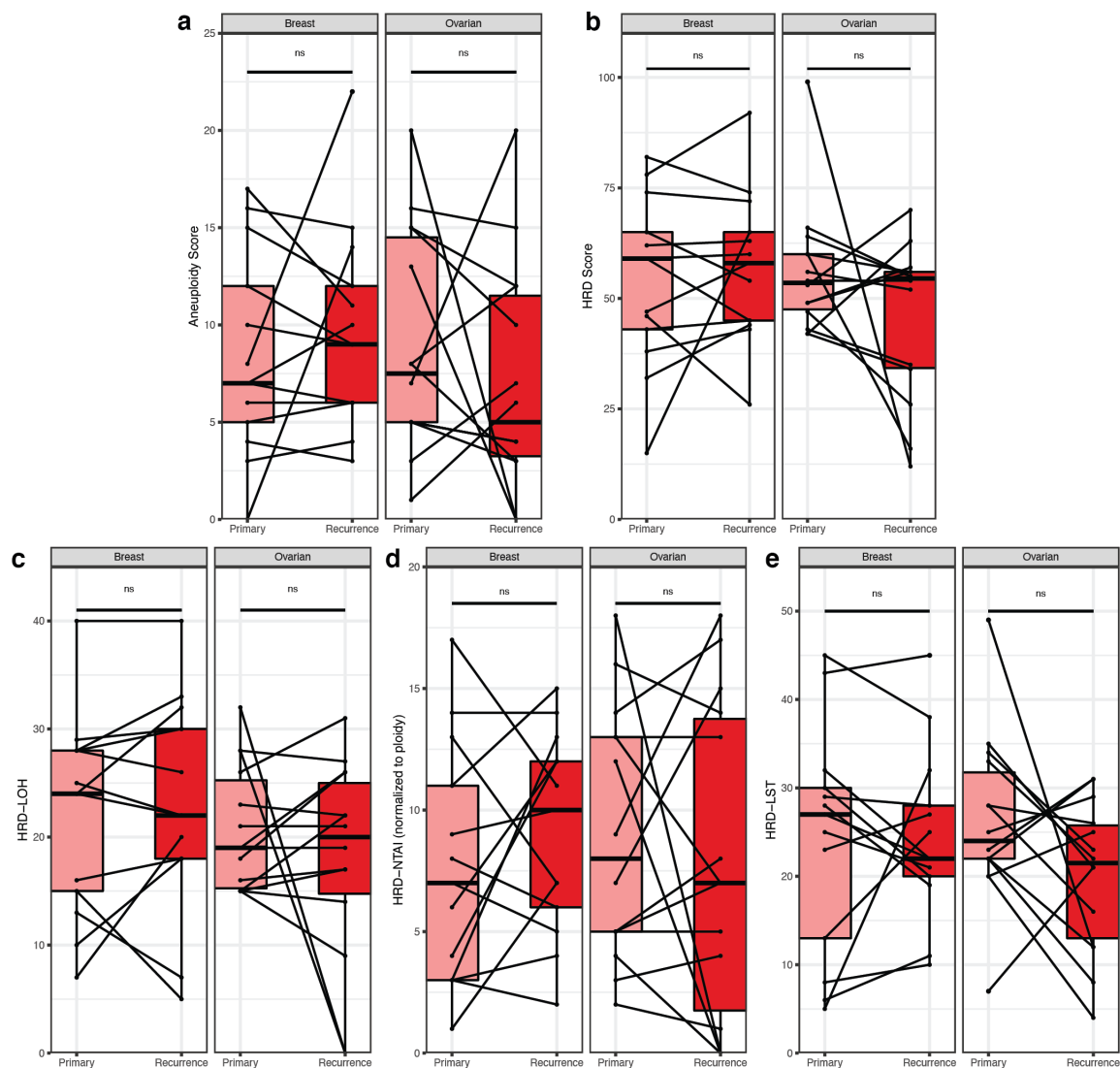


Figure 16: Comparisons of HRD and aneuploidy scores in primary/recurrent tumors. A. Comparison of aneuploidy score for primary/recurrent tumor pairs. B. Comparison of HRD score for primary/recurrent tumor pairs. C-E. Comparison of individual HRD metrics for primary/recurrent tumor pairs: loss of heterozygosity (C), non-telomeric allelic imbalance (D), and large scale state transitions (E). For patients with multiple recurrences, one recurrence was chosen at random for statistical comparisons in plots A-E. Pairwise differences were determined by two-sided Wilcoxon signed rank test ($\alpha=0.05$).

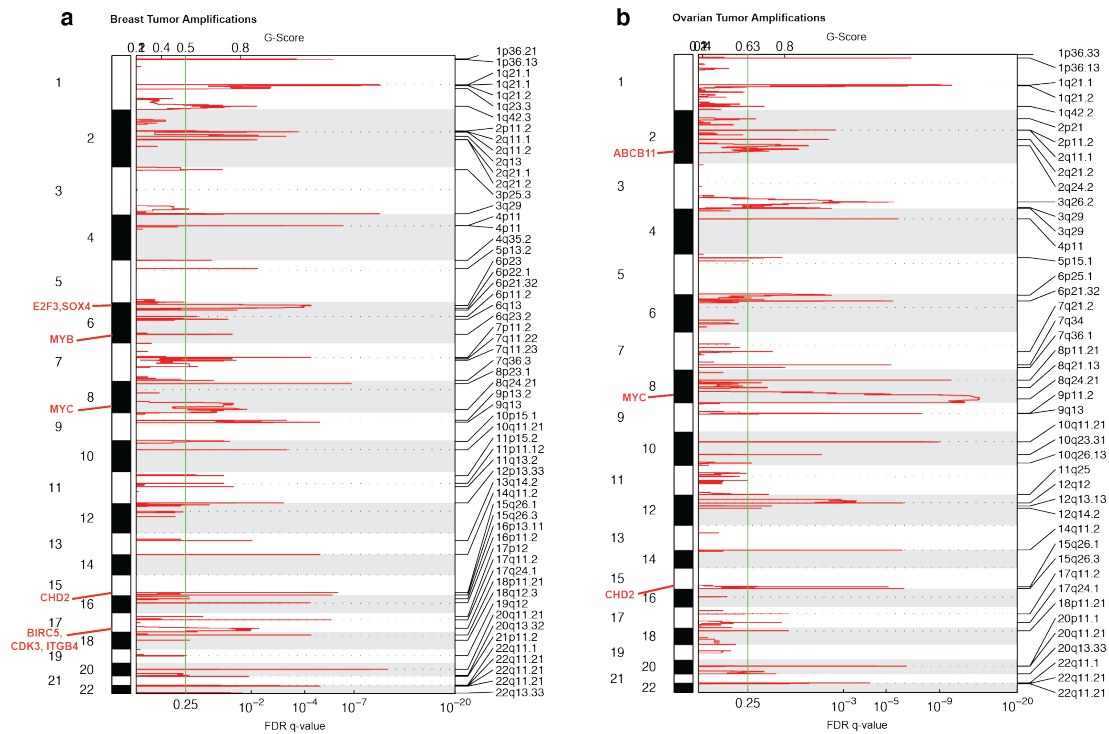


Figure 17: Copy number amplifications (GISTIC) in breast and ovarian tumors. A.

GISTIC qplot for 90% CI amplifications in (primary and recurrent) breast tumors. B.

GISTIC qplot for 90% CI amplifications in (primary and recurrent) ovarian tumors. All

highlighted genes in A and B have residual $q < 0.05$.

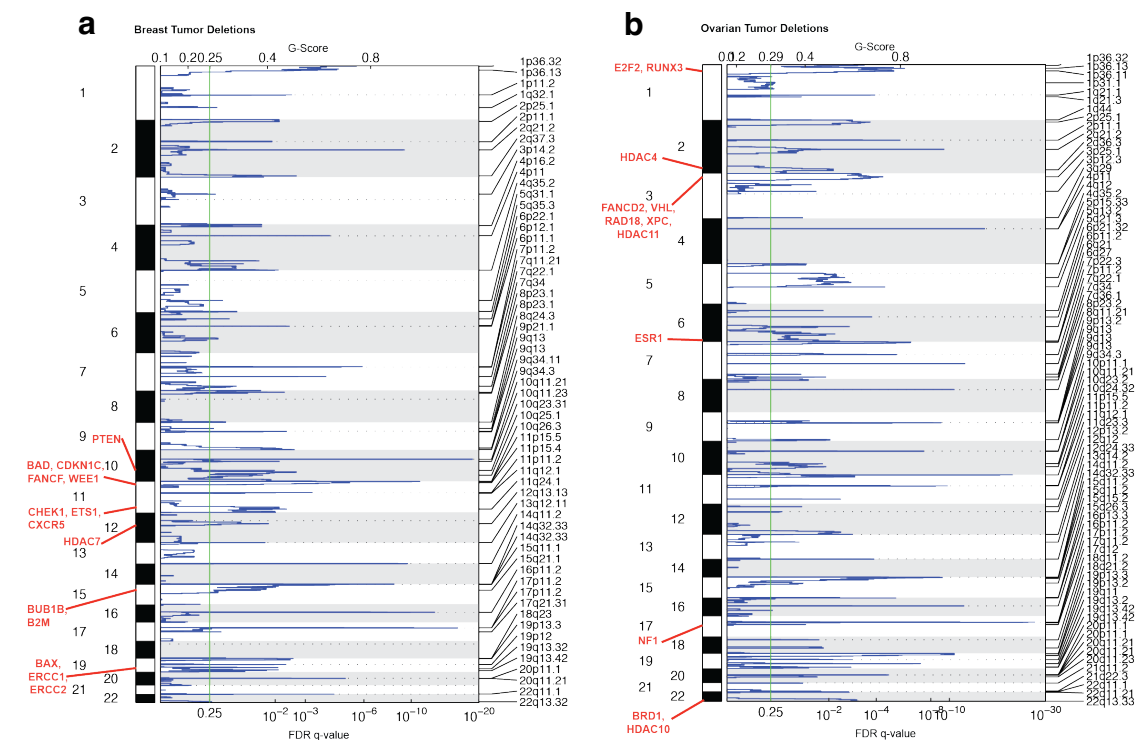


Figure 18: Copy number deletions (GISTIC) in breast and ovarian tumors. A.

GISTIC qplot for 90% CI deletions in (primary and recurrent) breast tumors. B. GISTIC qplot for 90% CI deletions in (primary and recurrent) ovarian tumors. All highlighted genes in A and B have residual $q < 0.05$.

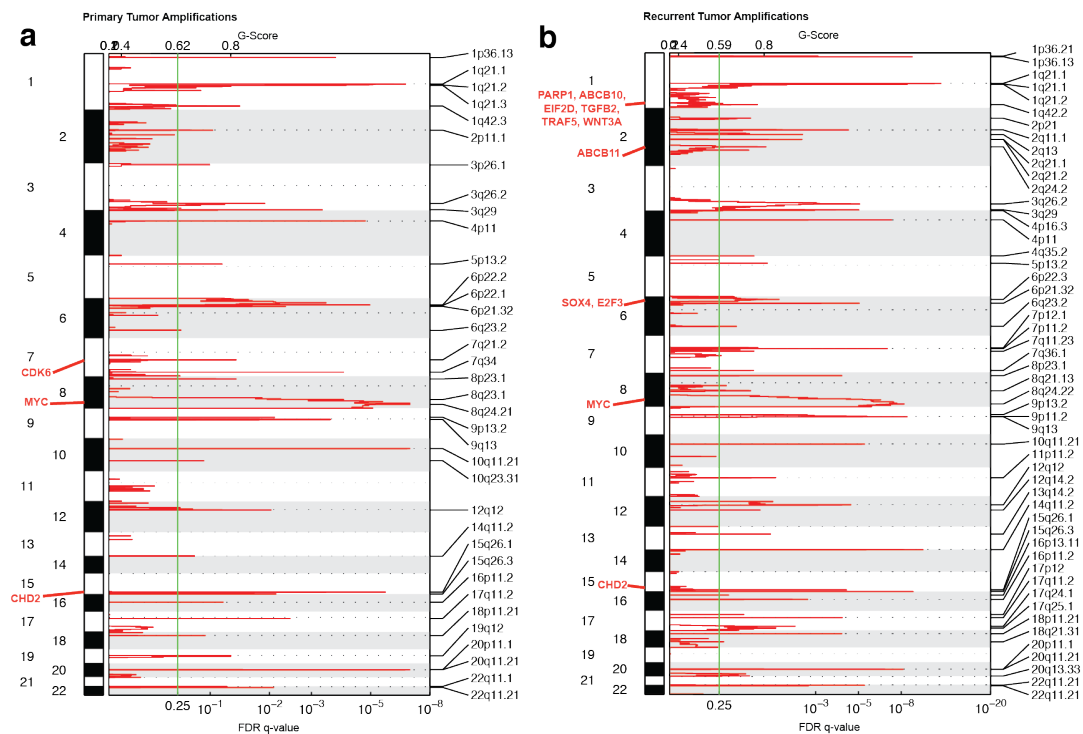


Figure 19: Copy number amplifications (GISTIC) in primary and recurrent tumors.

A. GISTIC qplot for 90% CI amplifications in primary breast and ovarian tumors. B.

GISTIC qplot for 90% CI amplifications in breast and ovarian recurrences. All highlighted genes in A and B have residual $q < 0.05$.

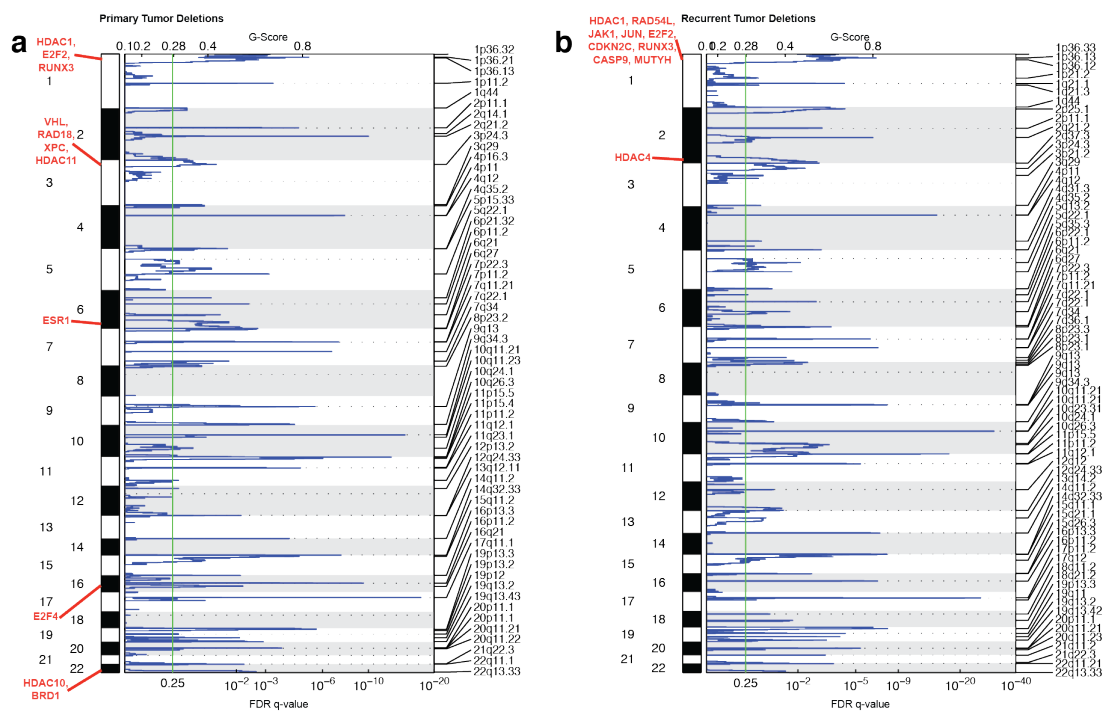


Figure 20: Copy number deletions (GISTIC) in primary and recurrent tumors. A. GISTIC qplot for 90% CI deletions in primary breast and ovarian tumors. **B.** GISTIC qplot for 90% CI deletions in breast and ovarian recurrences. All highlighted genes in A and B have residual $q < 0.05$.

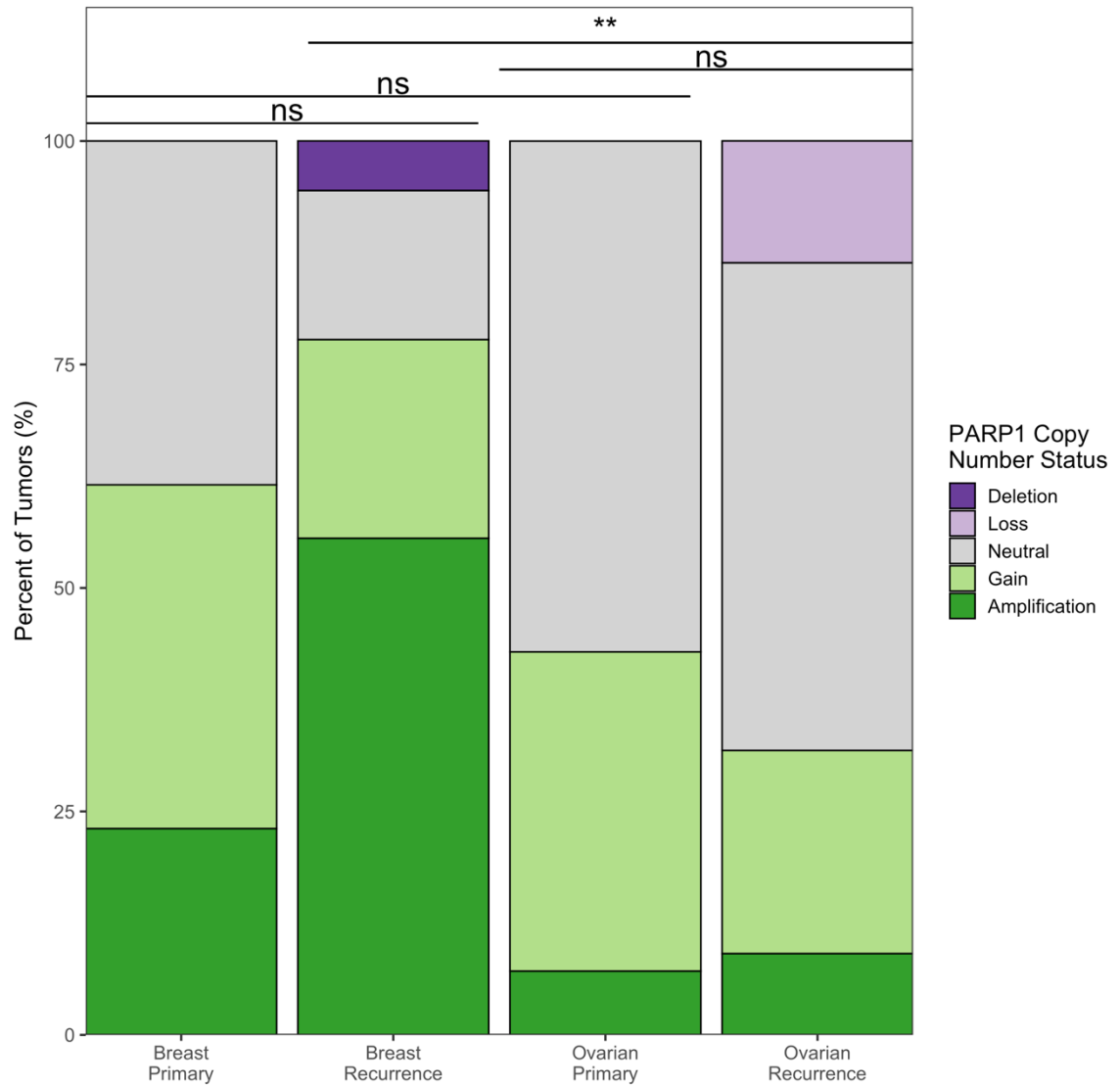


Figure 21: *PARP1* copy number variation in primary/recurrent cohort. A. *PARP1* copy number variation by tumor in primary/recurrent cohort. Groupwise differences in average copy number were determined by Kruskal-Wallis test, followed by Dunn's test with Bonferroni correction ($\alpha=0.05$, $**p<0.01$). Total copy number (Sequenza) was binned as follows: Deletion, CN=0; Loss, CN=1; Neutral, CN=2-3; Gain, CN=4-5; Amplification, CN>5.

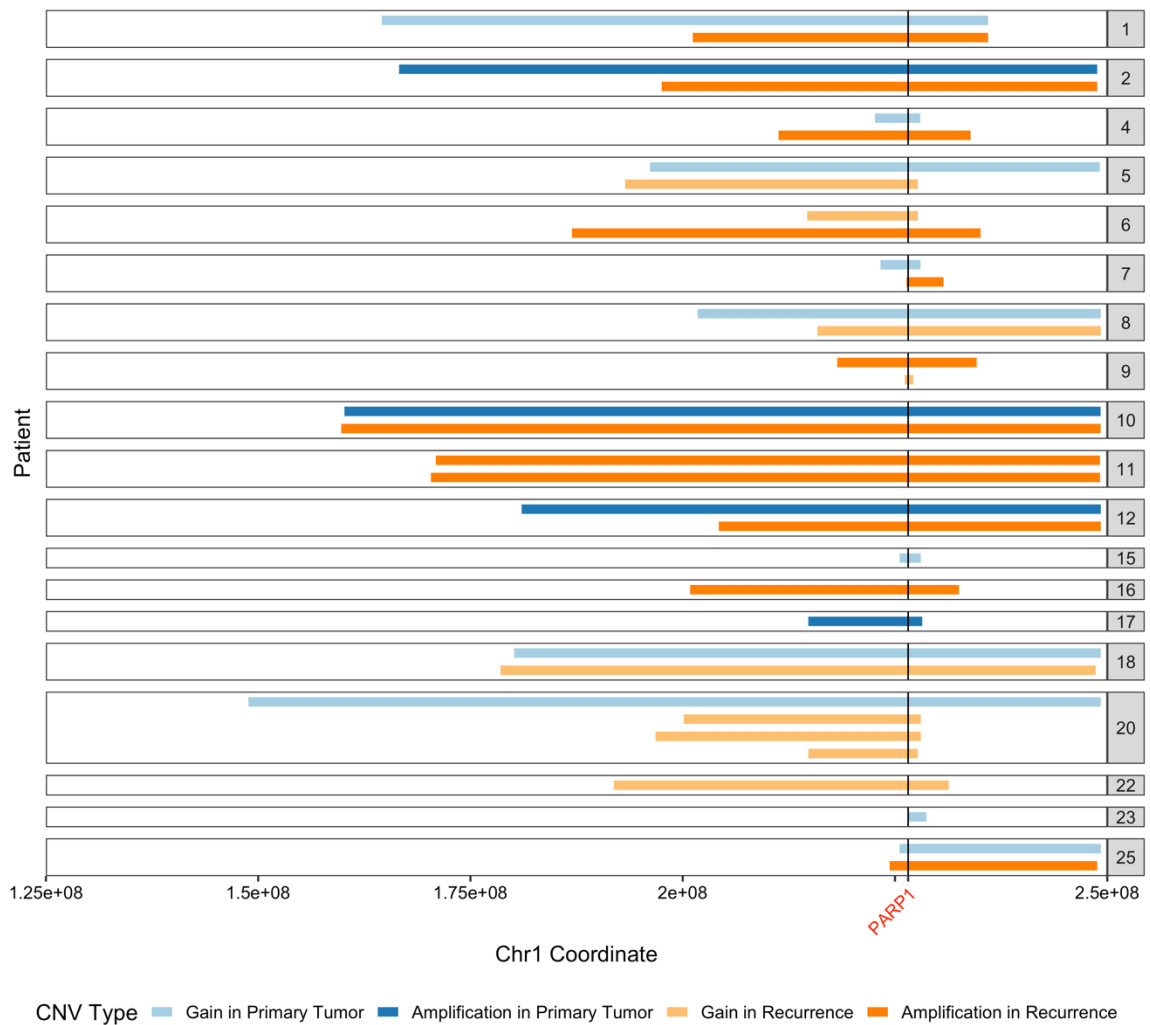


Figure 22: Segments of *PARP1* gains and amplifications by tumor. Segments of copy number gains and amplifications encompassing *PARP1*, by patient. Total copy number (Sequenza) was binned as follows: Deletion, CN=0; Loss, CN=1; Neutral, CN=2-3; Gain, CN=4-5; Amplification, CN>5.

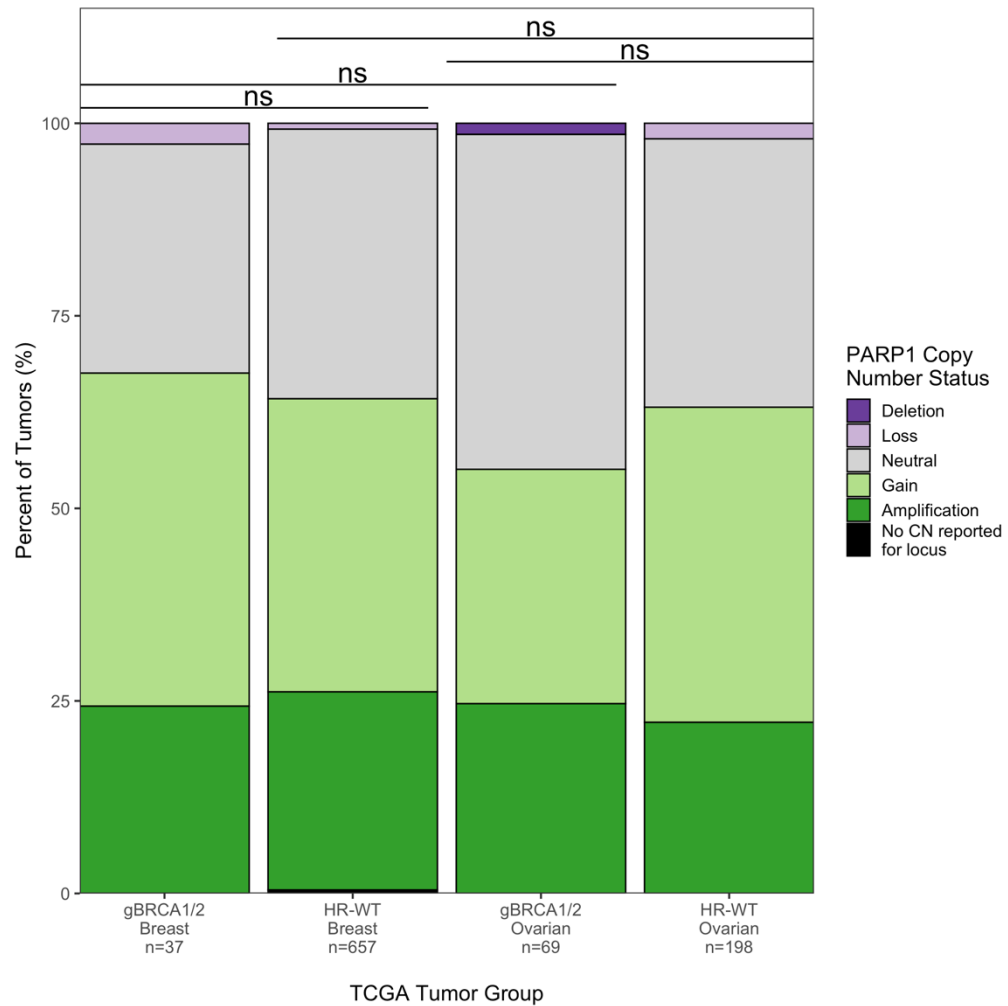


Figure 23: *PARP1* copy number variation in TCGA PanCancer cohorts. *PARP1* copy number by tumor group in TCGA breast and ovarian cohorts. Tumors were grouped into those with germline *BRCA1/2* mutations (*gBRCA1/2*) and those without germline or somatic loss of HR genes (HR-WT; see **Methods** and **Figure 11**). Groupwise differences in average copy number were determined by two-sided t-test ($\alpha=0.05$). Copy number (Sequenza) was binned as follows: Deletion, CN=0; Loss, CN=1; Neutral, CN=2-3; Gain, CN=4-5; Amplification, CN>5.

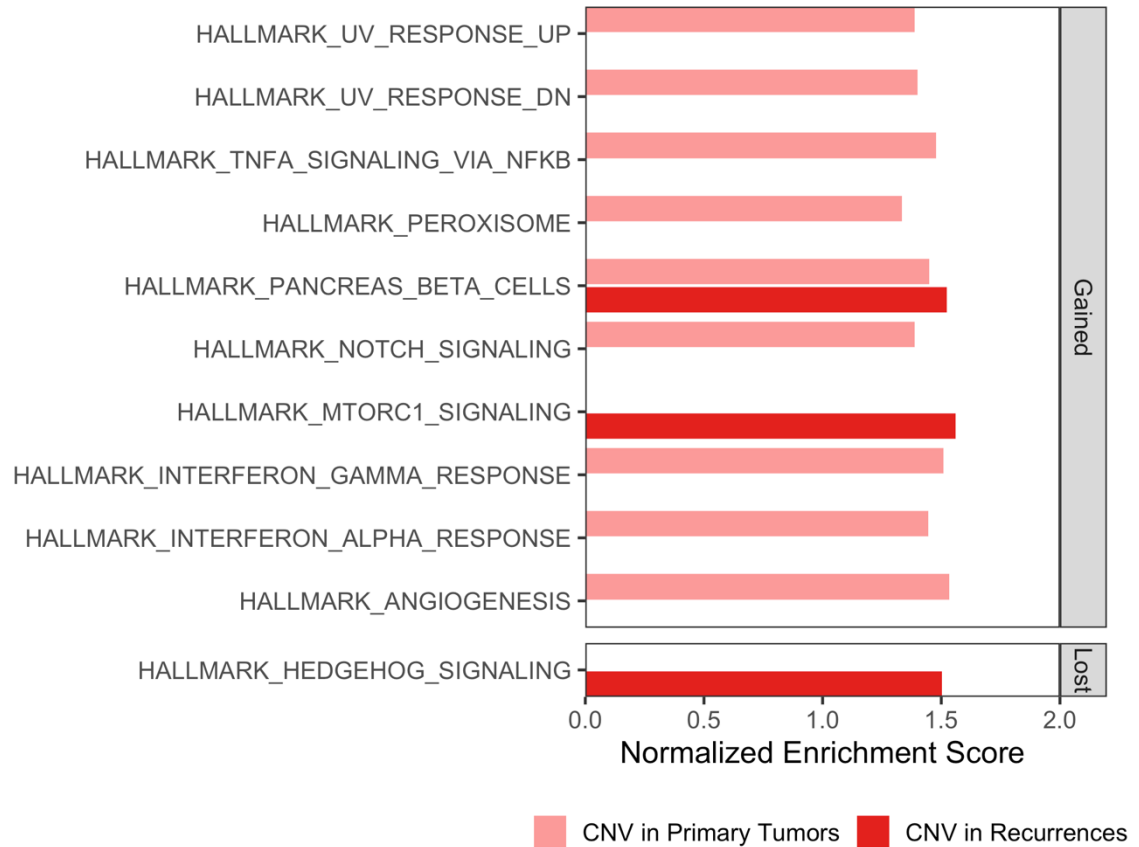


Figure 24: Pathway analysis of genes subject to copy number variation. Hallmark Gene Sets enriched in primary and recurrent tumor gains and amplifications ($CN \geq 4$, top) vs. losses and deletions ($CN \leq 1$, bottom). All gene sets displayed have FDR $q < 0.25$.

Discussion

This study is the largest integrated analysis of copy number variation, LOH status, HRD, and aneuploidy in *BRCA1/2* mutation-associated tumors. First, we found that seven pairs of primary and recurrent tumors demonstrated changes in allele-specific LOH status. Discordant LOH status was more common in breast cancer patients. We observed the highest frequency of nonLOH tumors (50% of tumors, 100% of patients) in *BRCA2* mutation-associated breast cancers, as has been previously reported³¹. We

observed a transition to allele-specific LOH in three breast cancers (one *BRCA1*, two *BRCA2*) upon recurrence, associated with increased HRD scores. Interestingly, two *BRCA2* mutation-associated breast tumors and two *BRCA1* mutation-associated ovarian tumors LOH reversal. This observation was corroborated by lower HRD scores in respective nonLOH recurrences. Prior studies have suggested LOH reversal occurs in a subset of *BRCA1* mutation-associated ovarian cancers after neoadjuvant platinum treatment, as seen here^{45–47}. We also identified LOH reversal in *BRCA2* mutation-associated breast cancers treated with (non-platinum) chemotherapy and radiation. We hypothesize that LOH reversal is mediated by clonal selection, in which rare nonLOH cells are present in the primary tumor, survive genotoxic treatments, and seed subsequent recurrences. Ultimately, our results suggest that some tumors evolve towards HR proficiency (as evidenced by nonLOH recurrences) while others evolve towards HR deficiency (as evidenced by recurrences with LOH).

Primary and recurrent tumors showed a number of differences in their CNVs. Breast and ovarian primary tumors had distinct copy number amplification and deletion profiles compared to breast and ovarian recurrences. Notably, recurrences of both tumor type had amplifications in ABC transporter *ABCB10*, suggesting a potential shared mechanism of chemoresistance¹²². Recurrences exclusively accumulated copy number deletions and LoF mutations in Hedgehog signaling genes, suggesting that loss of Hedgehog signaling could be a shared driver of both breast and ovarian tumor recurrence. Primary tumor gains spanned more pathways than recurrences, pointing to high NF- κ B, IFN- α , and IFN- γ signaling as potential drivers in this setting. Conversely, recurrences only had amplifications in one pathway (MTORC1 signaling), suggesting a different growth mechanism in recurrent disease.

Primary and recurrent tumors also had a number of similarities. In agreement with prior studies comparing *BRCA1/2* mutation-associated ovarian tumors pre- and post-platinums, HRD scores generally did not increase upon recurrence¹²³. Aneuploidy scores also did not increase. Taken together, these results suggest that *BRCA1/2* mutation-associated tumors accumulate the majority of their chromosomal abnormalities early in tumorigenesis, and that these do not significantly worsen upon recurrence. Another consistent feature between primary and recurrent tumors were *MYC* amplifications, which were significant for both groups. This finding agrees with prior reports of *MYC* amplifications in *BRCA1/2* mutation-associated primary tumors, but is the first report of *MYC* amplifications retained in recurrence^{124–126}.

Another common feature across the cohort was *PARP1* gains and amplifications, which were observed in all tumor groups. However, these gains were only statistically significant in recurrent tumors, possibly driven by the high frequency (78%) in breast cancer recurrences. We identified *PARP1* amplifications up to eight copies, centering on a minimal common region of chr1:226,474,131-227,148,258. This minimal region fully encompasses only three other genes (*ITPKB*, *PSEN2*, *STUM*) in addition to *PARP1*, suggesting that *PARP1* amplifications drove selection for amplifications in this region. In TCGA cohorts, we found a similar frequency of *PARP1* gains and amplifications as in the primary tumors we evaluated; and further, prevalence did not significantly differ between g*BRCA1/2* and sporadic HR-WT breast and ovarian tumors. Variation in analytic methods likely accounts for the slightly different prevalence of *PARP1* gains and amplifications reported in cBioPortal for TCGA PanCancer ovarian (53%) and breast (74%) cohorts¹²⁷.

As PARPi are increasingly employed to treat *BRCA1/2* mutation-associated

breast and ovarian tumors, *PARP1* amplifications and overexpression may have clinical relevance as a biomarker of therapeutic response or clinical outcomes in recurrences. Results from our cohort suggest that *PARP1* gains and amplifications are more prevalent than *PARP1* losses in *BRCA1/2* mutation-associated tumors. *PARP1* only underwent copy number loss in single *BRCA1* ovarian cancer patient and a single *BRCA2* breast cancer patient, neither of whom received PARPi. Further, while limited in sample size, our findings do not support the conclusion that *PARP1* loss is a common mechanism of resistance to PARPi.

We also investigated the effects of copy number variation on other genes related to platinum and PARPi resistance (*ABCB1*, *MAD2L2/REV7*, *TP53BP1*, and *CHD4*). Copy number gains and amplifications were noted for *ABCB1* in 44% of primary tumors and 50% of recurrences, with a high degree of conservation in primary/recurrent tumor pairs (92%). We hypothesize that gains in *ABCB1* (and *ABCB10*, noted above) could facilitate increased expression of the encoded efflux pumps and mediate chemoresistance for these tumors^{51,71,122}. We did not observe *CHD4* loss in any *BRCA2* mutation carriers treated with platinum or PARPi. We observed *TP53BP1* loss in four *BRCA1* mutation-associated ovarian recurrences after platinum. One of these recurrences was exposed to PARPi and also harbored the only *MAD2L2/REV7* loss observed in the cohort. Overall, concurrent *TP53BP1* and *REV7/MAD2L2* losses represent potential routes to PARPi resistance in Patient 19, making this the first observation of either mechanism in a human patient with a germline *BRCA1* mutation.

CHAPTER 5: DIFFERENTIAL GENE EXPRESSION AND GENE FUSIONS

Introduction

Since most prior studies lack a transcriptomics component, gene expression in *BRCA1/2* mutation-associated cancer is poorly characterized. Due to low quality and yields of FFPE-derived RNA, traditional bulk RNA-seq methods do not perform well on such specimens¹²⁸. The requirement for fresh frozen tumor specimens for RNA-seq has historically impeded longitudinal tumor analyses, since FFPE samples are more commonly banked. However, with the advent of hybridization capture-based RNA-seq (including Illumina's TruSeq RNA Exome, used here), it is now possible to faithfully sequence mRNA from FFPE specimens^{128–131}. RNA Exome specifically captures RNA from coding regions of the genome, similar to the approach used in WES^{129–131}. This approach also detects “rare” events (gene fusions, etc.) with fewer reads than traditional bulk RNA-seq¹³¹.

In this study, we leveraged RNA Exome (hereafter referred to as “RNA-seq”) to assess differential gene expression and gene fusions in primary tumors, recurrences, and normal tissue from *BRCA1/2* mutation carriers. Due to sample availability, the RNA-seq cohort contained fewer primary/recurrent tumor pairs than used for WES. To maximize power in spite of this limitation, we focused our transcriptomic analysis on groupwise comparisons. We sought to understand what genes and pathways were differentially expressed across the cohort of tumor and normal tissue specimens, with the goal of identifying novel drivers of *BRCA1/2* mutation-associated cancer. All but four tumors in the RNA-seq cohort received WES (**Fig. 2**), so we also utilized RNA-seq as orthogonal evidence for trends observed in the WES arm of the project. Lastly, we used

RNA-seq to look for gene fusions (including activating fusions of *ABCB1*), since these could also represent tumor drivers^{51,71}.

Methods

Bioinformatic analysis of RNA sequencing

Fastq files from RNA sequencing were aligned using STAR aligner (v2.7.2a) and gene annotations from the GENCODE Human Release 19 reference assembly

(https://www.gencodegenes.org/human/release_19.html)⁷⁹. Alignment and quantification

of transcripts was performed using StringTie (v2.1.3b) to generate abundance files in

Ballgown readable format¹³². We used Tximport (v1.16.1) to quantify transcript

abundance from t_data.ctab files, generating “lengthScaledTPM” counts by gene¹³³.

Next, we used EdgeR (v3.30.3) to generate, filter, and normalize counts per million

(cpm)¹³⁴. Briefly, gene-level transcripts per million were used to create three separate

DGELists: one for all samples, one for ovarian tumor and normal fallopian tube samples,

and one for breast tumor and normal breast samples. All DGELists were filtered to

include only genes for which transcripts were detected in ≥ 2 or ≥ 3 samples, depending

on the size of the smallest biological group. Filtered DGELists were then normalized

using the trimmed mean of M values (TMM) method¹³⁵. Filtered, normalized DGELists

were used to generate filtered, normalized cpm for each of the three comparisons. We

assessed the distribution of filtered, normalized cpm across the cohort to identify any

outlying samples (**Fig. 25**). Filtered, normalized cpm were used for all cohort-level and

tumor-specific RNA-seq analyses, except for detection of isoform switching (see Chapter

6).

Differential gene expression and functional enrichment analyses

Filtered, normalized $\log_2\text{cpm}$ were used to assess sample relatedness. We performed hierarchical clustering (maximum distance, average agglomeration) to generate dendrograms. Dendrograms and principal components analysis (PCA) plots were generated using the stats package in base R (v4.0.2). Between-group fold changes were computed as the difference in average $\log_2\text{cpm}$ for a given gene between groups. Using $\log_2\text{cpm}$ fold changes as input, we ran GSEA with Hallmark gene sets using clusterProfiler (v3.16.1) and msigdb (v7.2.1).

We next used the limma R package (v3.44.3) to model mean-variance separately for all-tumor, ovarian-only, and breast-only DGELists. Bayesian statistics were extracted from the resulting linear models to generate adjusted p-values (Benjamini-Hochberg method) and $\log_2(\text{fold change})$ by gene for each comparison. Significantly altered genes were extracted using decideTests in limma (global method, $\text{adj.p} < 0.05$, $|\log_2(\text{fold change})| > 1$).

Significantly altered genes for all-tumor comparisons were clustered into two gene modules (see **Supplementary File 9**). Hierarchical clustering of genes was based on Pearson coefficients from pairwise correlations (complete method). Hierarchical clustering of samples was based on Spearman coefficients from pairwise correlations (complete method). A heatmap was generated from the resulting dendrograms using the gplots package (v3.1.0). Gene Ontology was performed and plotted for each gene module separately, using gProfiler2 (v0.2.0). Enrichment of transcription factor binding motifs was performed using oPOSSUM (v3.0) Single Site Analysis (<http://opossum.cisreg.ca/oPOSSUM3/>)¹³⁶. OPOSSUM analysis was conducted against JASPAR vertebrate core profiles with cutoffs for conservation score > 0.40 , oPOSSUM Fisher score ≥ 7 , and oPOSSUM Z-score ≥ 10 .

Detection of gene fusions from RNA sequencing

Gene fusions were identified using STAR-Fusion (v1.7.0) and plots were generated using FusionInspector (v2.1.0)¹³⁷. We limited our analysis of gene fusions to those with ≥ 5 junction-spanning reads.

Immunohistochemistry (IHC) analysis of PARP1

Twenty-three primary and recurrent *BRCA1/2* mutation-associated tumors (a subset of the primary/recurrent cohort used throughout the study) were assembled into three tissue microarrays (TMAs). Each tissue microarray contained 4-11 1mm cores per tumor to account for tumor heterogeneity. TMA blocks were constructed, cut, stained, and University of Pennsylvania. Immunohistochemistry for PARP1 was performed using standard laboratory protocols and the PARP (46D11) Rabbit mAb (Cell Signaling Technology, catalog #9532). Stained slides were imaged at 20x resolution. PARP1 nuclear positivity was quantified for each core using H-score, as determined by Dr. Anupma Nayak. Hematoxylin and eosin (H&E) staining was used to assess tumor content in parallel; cores with inadequate tumor per H&E were excluded by Dr. Nayak. We calculated an average PARP1 H-score for each tumor for downstream analyses (range 3-10 evaluable cores per tumor).

Statistical analysis

Groupwise differences in *PARP1* mRNA expression and PARP1 H-score were assessed by Kruskal-Wallis test, with significant results followed by Dunn's test with Bonferroni correction ($\alpha=0.05$).

Results

Global transcriptomic analysis of primary and recurrent tumors

We performed RNA sequencing (RNA-seq) on 50 primary and recurrent tumors in the cohort, adding in recurrences from four additional patients (subjects 28-31; **Supplementary File 1**). We also sequenced 12 normal breast and fallopian tube samples from prophylactic surgeries in *BRCA1/2* mutation carriers with no history of cancer (subjects 32-43, **Fig. 26, Supplementary File 1**). First, we used hierarchical clustering to assess sample relatedness across the cohort (**Fig. 27**). In the resulting dendrogram, many paired primary and recurrent tumors were not each other's closest relatives (11/19 tumor pairs). In general, ER- breast tumors clustered closer to ovarian tumors than to ER+ breast tumors. ER+ breast tumors clustered closest to normal samples. Four ER+ ovarian recurrences came from Patient 21, whose tumors were evaluated for hormonal status because she was initially misdiagnosed with breast cancer. These ER+ ovarian recurrences were most closely related to the ER+ breast tumors. Thus, sample relatedness appeared to be most informed by ER status, rather than tissue of origin.

We further assessed sample clustering using principal components (**Fig. 28**). We did not observe any clustering associated with tumor origin (breast or ovarian), tumor type (primary or recurrent), or germline mutation (*BRCA1* vs. *BRCA2*). Ovarian recurrences formed the most homogenous group on the plot, likely due to similarities in treatment (21/23 received platinum/taxane therapy) and anatomical location (16/23 were intra-abdominal).

To identify groupwise differences in gene expression by tumor type, we

combined primary and recurrent tumors and compared to tissue-matched normal controls. Gene expression was highly similar between primary and recurrent tumors of the same origin; the average \log_2 (counts per million) for individual genes was highly correlated between breast ($R^2=0.972$) and ovarian ($R^2=0.965$) primary and recurrent tumors. One significantly overexpressed gene in ovarian cancers was *ASPM*, assembly factor for spindle microtubules, which also was highly expressed in breast tumors (**Fig. 29, 30**). *ASPM* expression has been correlated with ovarian cancer staging, as well as stemness and prognosis in other solid tumors^{138–140}. Among the genes with the highest log-fold change in ovarian tumors was *BIRC5*, which inhibits apoptosis and has been implicated in progression of many cancers (**Fig. 29b, 30b**)¹⁴¹. The second most significantly under-expressed gene in ovarian cancer was *RERGL*, a potential *BRCA1* transcriptional target whose expression is associated with decreased survival in colorectal cancer¹⁴².

As the average \log_2 (counts per million) for individual genes were similar between breast and ovarian tumors overall ($R^2=0.934$), we combined all *BRCA1/2* mutation-associated tumors to identify the major transcriptomic differences compared to normal tissues. Using GSEA, we found enrichment of 25 Hallmark gene sets (all adj. $p<0.01$, **Fig. 31, Supplementary File 8**). The identified gene sets included several pathways also enriched for copy number gains (**Fig. 24**): $IFN\alpha$ and $IFN\gamma$ signaling, MTORC1 signaling, response to UV light, and $TNF\alpha$ signaling via $NF-\kappa B$. Increased expression of these gene sets was also identified in GSEA stratified by tumor type (**Fig. 32**).

To identify transcriptomic trends driving tumorigenesis, we next interrogated genes that were significantly different ($|\log_2(\text{fold change})|>1$ and adj. $p<0.05$) in tumor vs. normal tissue. Restricting our analysis to these genes, hierarchical clustering identified

two modules: 1) 4715 genes with low expression in tumors and 2) 2492 genes with high expression in tumors (**Fig. 33, Supplementary File 9**). Under-expressed genes (module 1) were enriched for transcription factor binding motifs for lineage-related transcription factors (Forkhead and Homeobox families) and myogenic regulators *MEF2A* and *SRF* (all Fisher score ≥ 7 , Z-score ≥ 10). Using Gene Ontology (GO), we evaluated the genes in module 1. Among the top hits for transcription factor binding sites lost in tumors, we identified TEF-3, a transcription factor for Hippo signaling ($p=1.8 \times 10^{-30}$); and PARP, a predicted binding motif for PARP1 ($p=6.6 \times 10^{-29}$) (**Fig. 34a**)¹⁴³.

We also used GO to establish trends for genes with increased expression in tumors (module 2) and found enrichment of terms for mitosis and cell cycle (**Fig. 34b**). In agreement with GSEA results, we also found enrichment of several immune activation terms in tumors (all $p < 1 \times 10^{-7}$). GO results also indicated that tumors have high expression of hsa-miR-193b-3p, a demonstrated tumor suppressor in ovarian cancer and TNBC (top miRNA result, $p=3.2 \times 10^{-23}$)^{144,145}. This miRNA was shown to impede HR through direct targeting of *BRCA1*, *BRCA2*, and *RAD51* mRNAs in breast and ovarian cancer cell lines¹⁴⁶. Therefore, even if acting as a tumor suppressor, miR-193-3p could also contribute to HRD in *BRCA1/2* mutation-associated cancers.

To identify determinants of recurrence in these cancers, we next compared the transcriptomes of primary and recurrent tumors stratified by tumor type. Gene expression was highly similar between groups within each tumor type, such that no genes met our criteria for significantly altered expression ($|\log_2(\text{fold change})| > 1$ and adj. $p < 0.05$). Accordingly, we limited our primary/recurrent gene expression analyses to GSEA, which uses a ranking statistic approach optimized to detect small differences between groups. We found that breast primary tumors were enriched over breast

recurrences for KRAS and MTORC1 signaling, UV response, and several immune gene sets (**Fig. 35a**). However, ovarian recurrences demonstrated almost the opposite pattern, with enrichment over primary tumors for 32 Hallmark gene sets, encompassing genes for myogenesis, KRAS and MTORC1 signaling, UV response, inflammatory response, TNF α signaling via NF- κ B, IL6/JAK/STAT3 signaling, IL2/STAT5 signaling, and apoptosis (**Fig. 35b**). These results point to potential differences in signaling between primary and recurrent tumor transcriptomes, varying by tumor type.

We found that *PARP1* expression was significantly increased in breast and ovarian recurrences compared to normal tissue ($\log_2FC=1.62$, adj. $p=2.52\times10^{-6}$ and $\log_2FC=9.44$, adj. $p=1.58\times10^{-6}$) (**Fig. 29**). *PARP1* expression was also increased in breast primary tumors ($\log_2FC=1.43$, adj. $p=3.26\times10^{-5}$) (**Fig. 30a**). Primary and recurrent ovarian tumors both showed increased expression of *MYBL2*, for which *PARP1* is a putative coactivator (**Fig. 29b, 30b**)¹⁴⁷. Primary and recurrent breast tumors showed increased expression of *MKI67* (Ki67), expression of which has been positively correlated with *PARP1* expression in other breast tumor cohorts (**Fig. 29a, 30a**)¹⁴⁸. Overall, *PARP1* expression in all tumors exceeded normal tissue samples, an effect that did not depend on the presence of a *PARP1* gain or amplification ($p=3.17\times10^{-7}$ and $p=5.01\times10^{-5}$ for tumors with and without *PARP1* gain or amplification, respectively, compared to normal tissue; **Fig. 36**).

Next, we sought to assess whether *PARP1* copy number gains and mRNA expression translated to high *PARP1* protein expression in this sample set. We assembled tissue microarrays (TMAs) containing 23 primary and recurrent *BRCA1/2* mutation-associated tumors from this cohort (**Supplementary File 1**). Each TMA contained cores from 4-11 sites in each tumor to account for spatial heterogeneity. We

performed immunohistochemistry (IHC) for PARP1, then quantified PARP1 nuclear positivity for each tumor by average H-score (on a 0-300 scale). We found that PARP1 protein was highly expressed across the cohort of tumors (mean H-score 252.5, median H-score 276.4) with no significant differences by tumor type (**Fig. 37a**). As in the RNA-seq results, high PARP1 expression did not depend on the presence of a PARP1 gain or amplification (**Fig. 37b**). However, the only tumor with a *PARP1* copy number loss had a relatively low H-score of 112. Tumors with *PARP1* gains and amplifications also had a smaller distribution of H-score than tumors that were copy neutral for *PARP1* ($\sigma=35.4$ vs. $\sigma=42.7$, respectively), suggesting that a difference might be seen between these groups with a larger sample set. Overall, our results from RNA-seq and IHC support the hypothesis that PARP1 is highly expressed in *BRCA1/2* mutation-associated breast and ovarian cancer, including those without *PARP1* copy number gains.

Expression of genes associated with platinum and PARPi resistance

In addition to *PARP1*, we assessed groupwise differences in expression of genes reported to mediate platinum and PARPi resistance (*ABCB1*, *CHD4*, *TP53BP1*, *MAD2L2/REV7*). In comparisons of breast or ovarian primary or recurrent tumor RNA vs. tissue-matched normal RNA, we did not identify any significant differences in expression of *CHD4*, *TP53BP1*, or *MAD2L2/REV7*. *ABCB1* expression was significantly lower in all four tumor vs. normal comparisons (all logFC < -2.75, adj.p < 0.01).

Analysis of gene fusions in primary and recurrent tumors

Lastly, we investigated whether gene fusions could be potential drivers of *BRCA1/2* mutation-associated tumors. We limited our analysis to fusions with ≥ 5 junction spanning reads (range 5-228 reads) and focused on the most common

biologically relevant gene fusions found in ≥ 3 patients. Six of 54 (11%) tumors had fusions involving genes for immunoglobulin light and/or heavy chains (IGL@, IGH@) (**Fig. 38; Supplementary File 10**). Fusions of *MALAT1*, a lncRNA associated with poor survival in metastasis in several tumor types, were recently observed at high frequency (13/18, 72%) in ovarian cancer and TNBC patients after PARPi treatment^{149,150}. We found *MALAT1* fusions in one breast and two ovarian (3/54, 5.5%) tumors. We did not identify any fusions involving ABC transporter gene *ABCB1*, which have been previously associated with multidrug resistance^{51,71}.

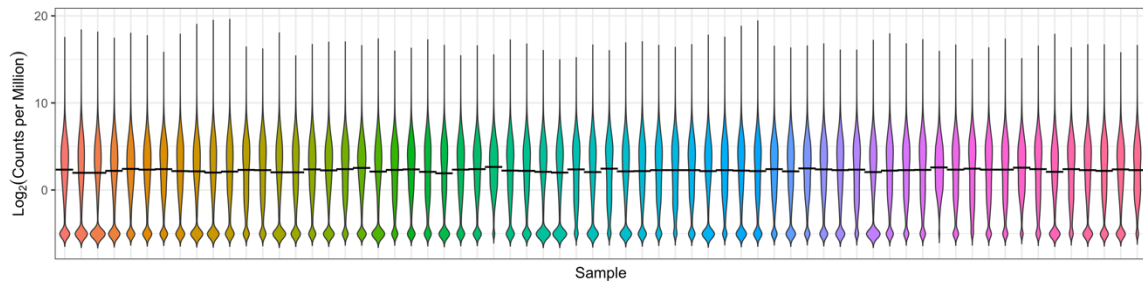


Figure 25: Distribution of $\text{Log}_2(\text{counts per million})$ across cohort of RNA-seq samples. Counts depicted are post-filtering and post-TMM normalization.

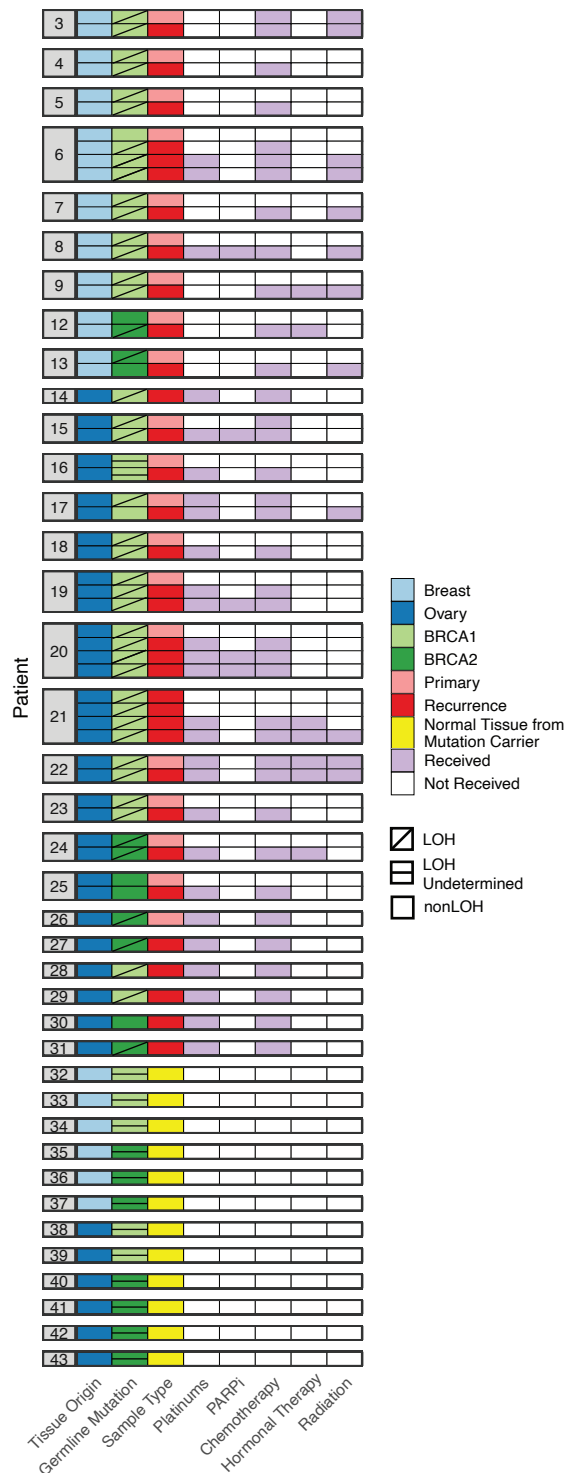


Figure 26: Cohort of 66 samples used

for RNA sequencing. Tumors are

displayed in chronological order by

patient, with the primary tumor at the top

and latest recurrence at the bottom.

Patients 28-31 had recurrent tumors only

(no primary tumor). Patients 32-43 were

BRCA1/2 mutation carriers with no prior

history of cancer or cancer treatment; their

samples are normal tissue from breast

and fallopian tube, and LOH status is

undetermined due to lack of whole exome

sequencing.

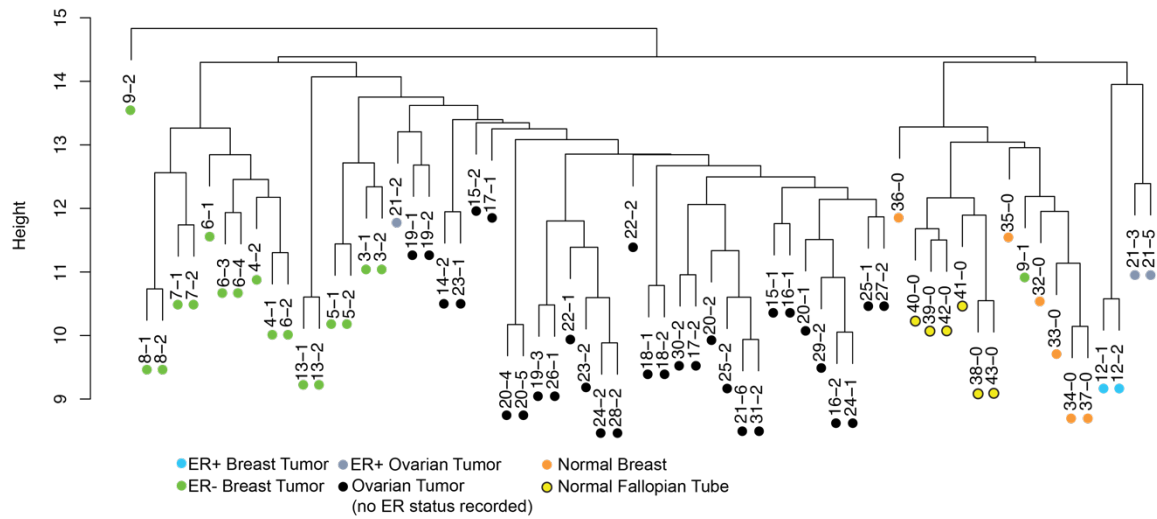


Figure 27: Dendrogram from agglomerative hierarchical clustering of RNA sequencing cohort. Samples are color coded by sample type and estrogen receptor (ER) status.

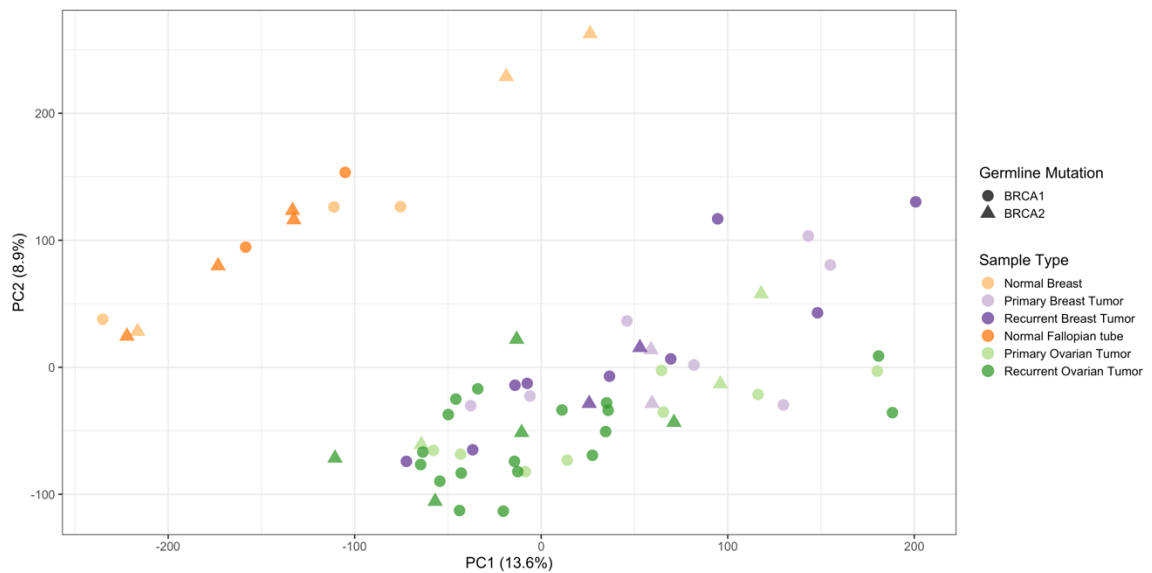


Figure 28: Principal components analysis of RNA-seq cohort. Samples are coded by sample type and germline mutation gene.

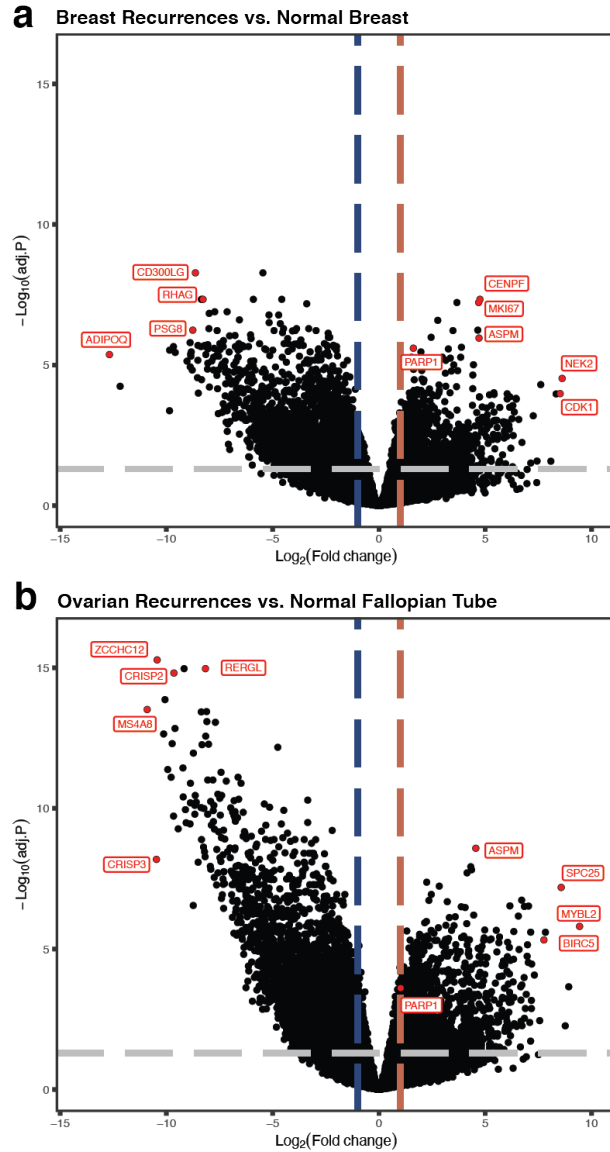


Figure 29: Differential gene expression in recurrent tumors vs. normal tissue. A. Differential gene expression in breast tumor recurrences vs. normal breast tissue from *BRCA1/2* mutation carriers. **B.** Differential gene expression in ovarian tumor recurrences vs. normal fallopian tube tissue from *BRCA1/2* mutation carriers. For A and B, a positive Log₂(fold change) indicates genes with increased expression in recurrent tumors. Adjusted p values were computed based on linear modeling of mean-variance trends (limma).

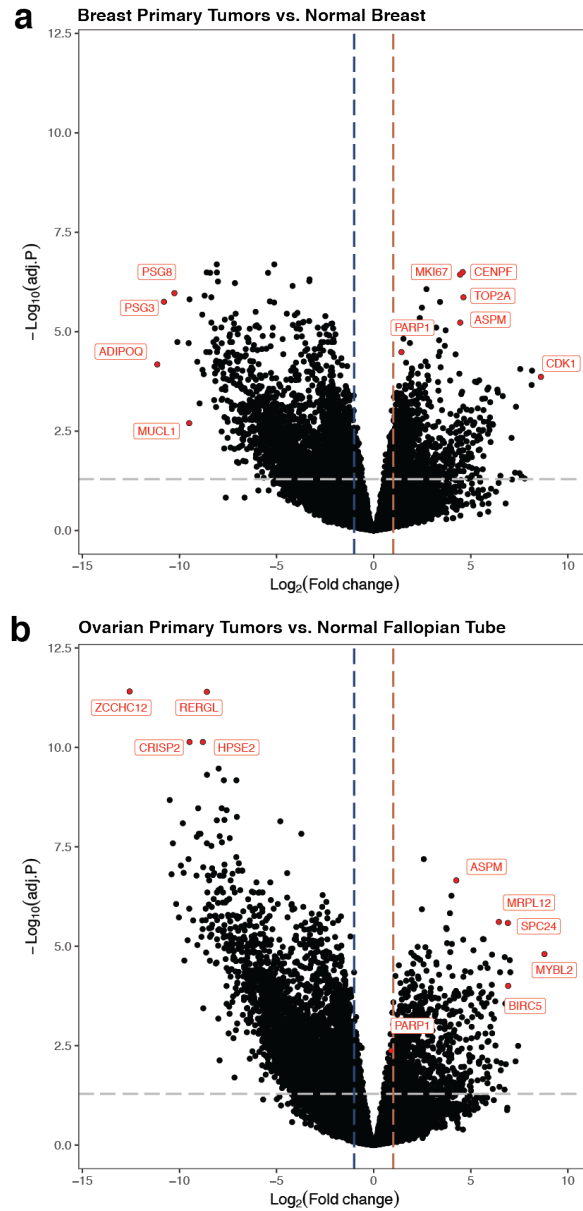


Figure 30: Differential gene expression in primary tumors vs. normal tissue. A.

Differential gene expression in primary breast tumors vs. normal breast tissue from *BRCA1/2* mutation carriers. B. Differential gene expression in primary ovarian tumors vs. normal fallopian tube tissue from *BRCA1/2* mutation carriers. For A and B, a positive $\text{Log}_2(\text{fold change})$ indicates genes with increased expression in primary tumors.

Adjusted p values were computed based on linear modeling of mean-variance trends (limma).

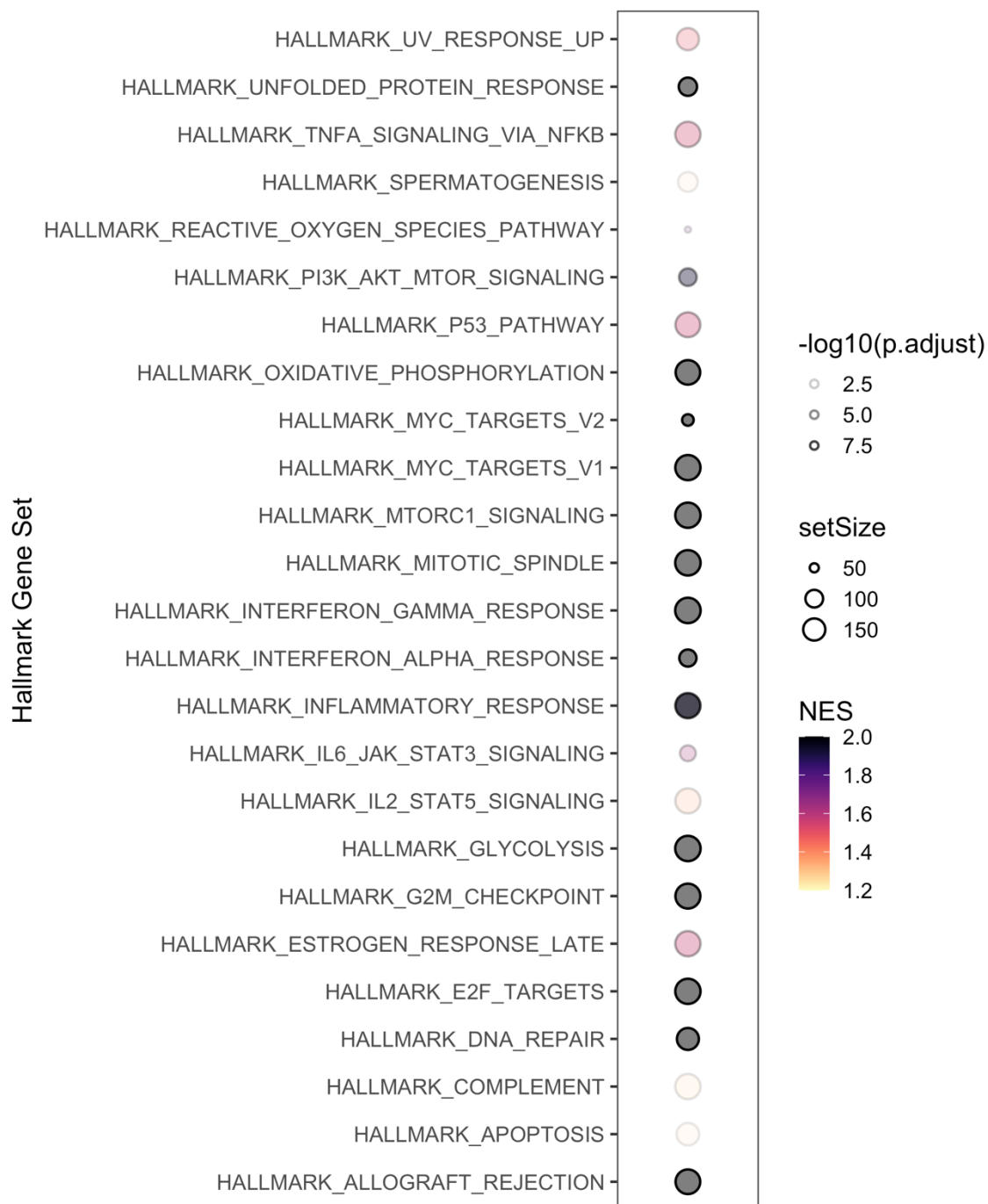


Figure 31: Pathway analysis of over-expressed genes in tumors. Hallmark Gene Sets enriched in genes with increased expression in primary and recurrent breast and ovarian tumors compared to normal breast and fallopian tube tissue (all adj.p<0.05).

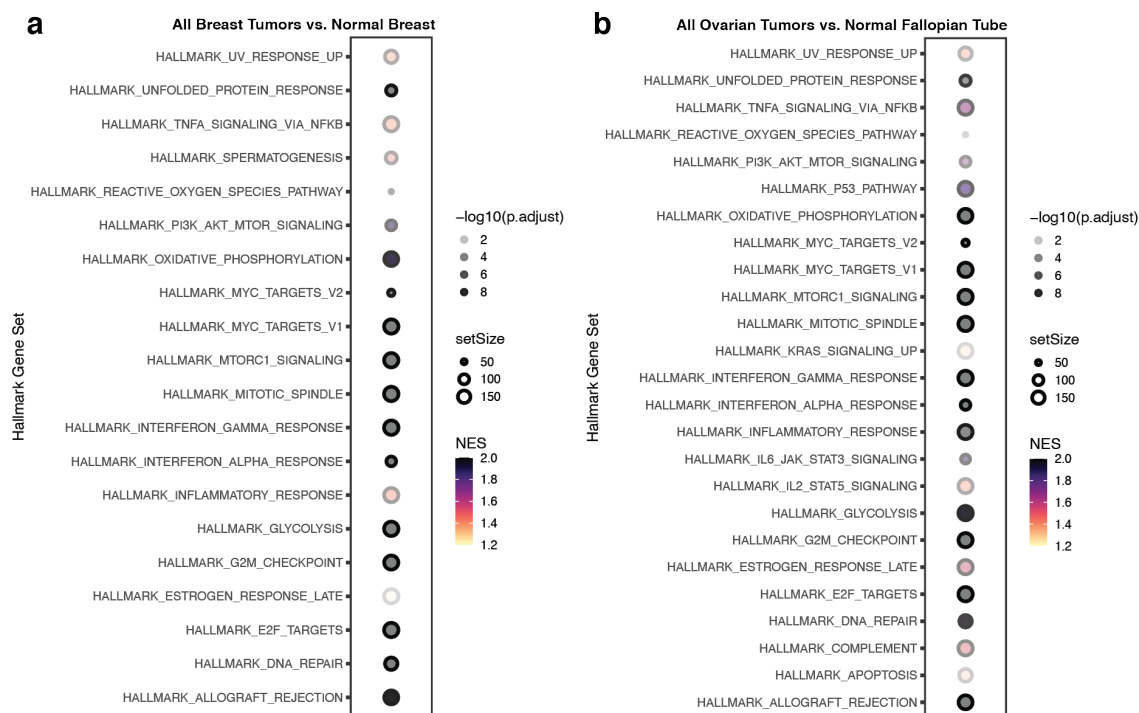


Figure 32: Pathway analysis of over-expressed genes in tumors by tumor type. A.

Hallmark Gene Sets enriched in genes with increased expression in primary and recurrent breast tumors compared to normal breast tissue from *BRCA1/2* mutation carriers. B. Hallmark Gene Sets enriched in genes with increased expression in primary and recurrent ovarian tumors compared to normal fallopian tube from *BRCA1/2* mutation carriers. For A and B, all gene sets had adj. $p < 0.05$.



Figure 33: Identification and functional enrichment of gene modules based on tumor vs. normal expression differences. Heatmap of hierarchical clustering to identify modules within set of differentially expressed genes with $\text{adj.p} < 0.05$ and $|\log_2\text{FC}| > 1$ for tumor vs. normal comparisons. Module of genes with decreased expression in tumor compared to normal tissue (coded yellow on heatmap) were enriched for transcription factor motifs as indicated (all Z-score ≥ 10 , Fisher score ≥ 7 for motif enrichment).

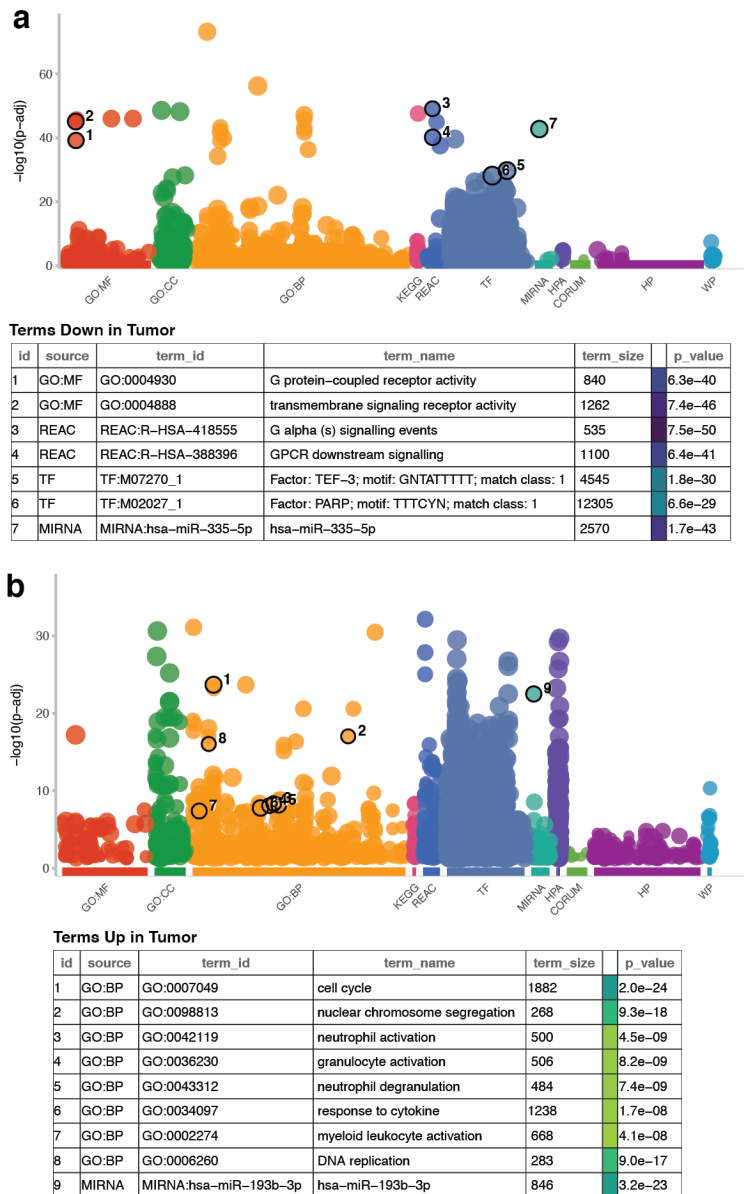


Figure 34: Functional enrichment of gene modules. A. Gene ontology terms enriched in module 1, genes with decreased expression in tumor vs. normal comparisons (coded yellow in **Fig. 33**). B. Gene ontology terms enriched in module 2, genes with increased expression in tumor vs. normal comparisons (coded blue in **Fig. 33**). For A and B, terms of particular biological relevance are highlighted in the table below each plot.

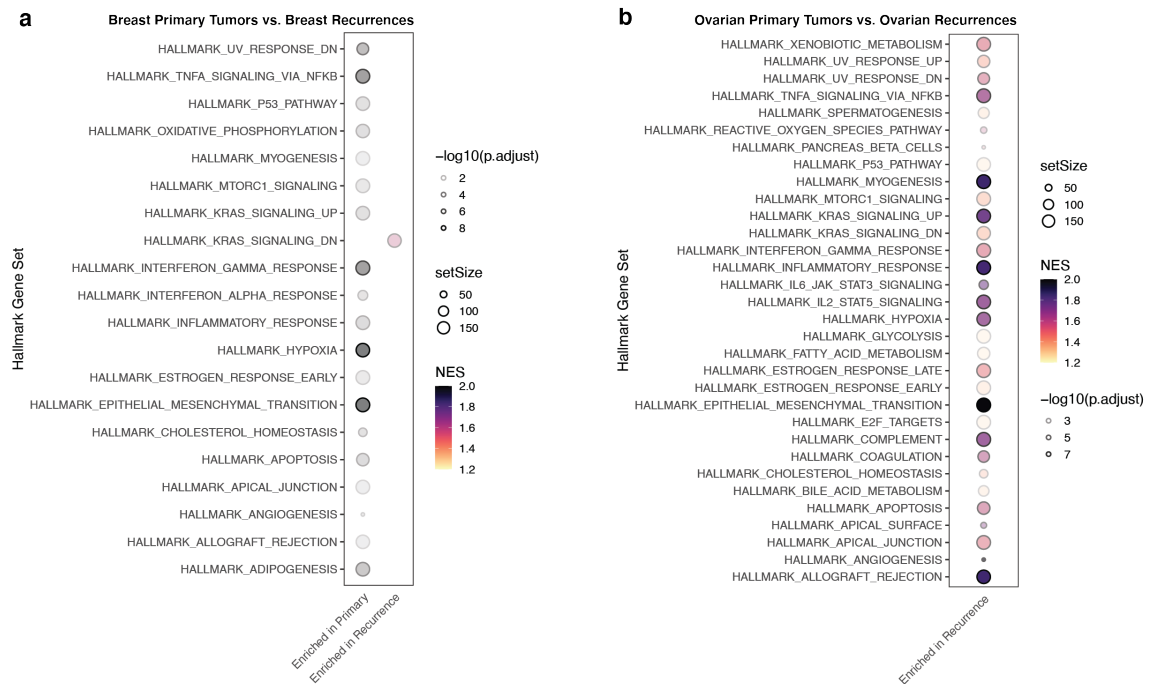


Figure 35: Gene set enrichment analysis in primary vs. recurrent tumors. A.

Hallmark Gene Set enrichment in differentially expressed genes from primary vs.

recurrent breast tumors. B. Hallmark Gene Set enrichment in differentially expressed

genes from primary vs. recurrent ovarian tumors. For A and B, all gene sets have adj.

$p < 0.05$.

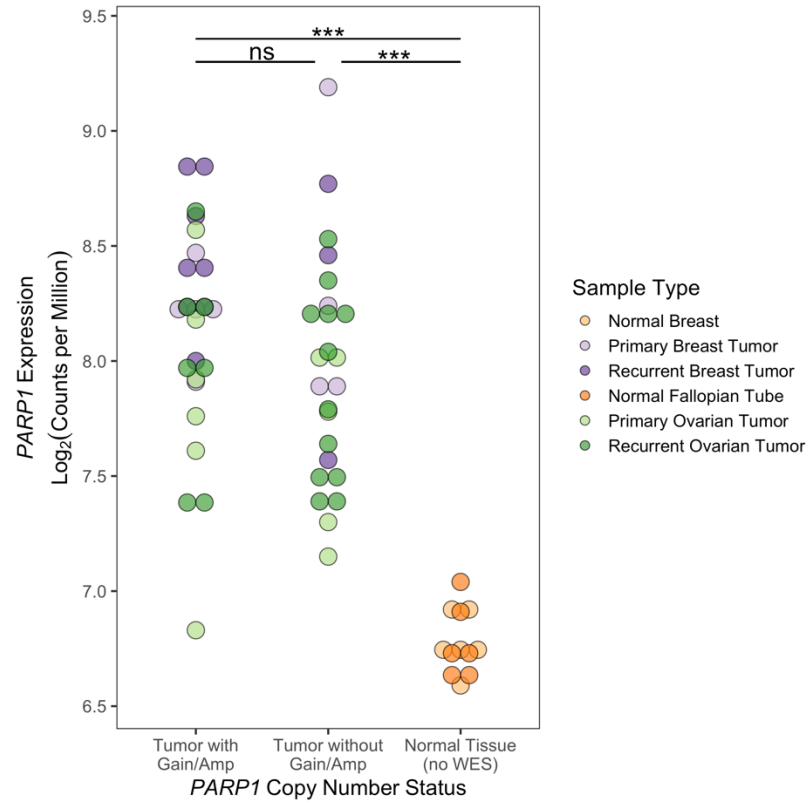


Figure 36: *PARP1* mRNA expression in primary and recurrent tumors. *PARP1* copy number status and expression across all tumors with whole exome and RNA-sequencing, and all normal samples with RNA sequencing. Groupwise differences in average copy number were determined by Kruskal-Wallis test, followed by Dunn's test with Bonferroni correction ($\alpha=0.05$, *** $p<0.0001$).

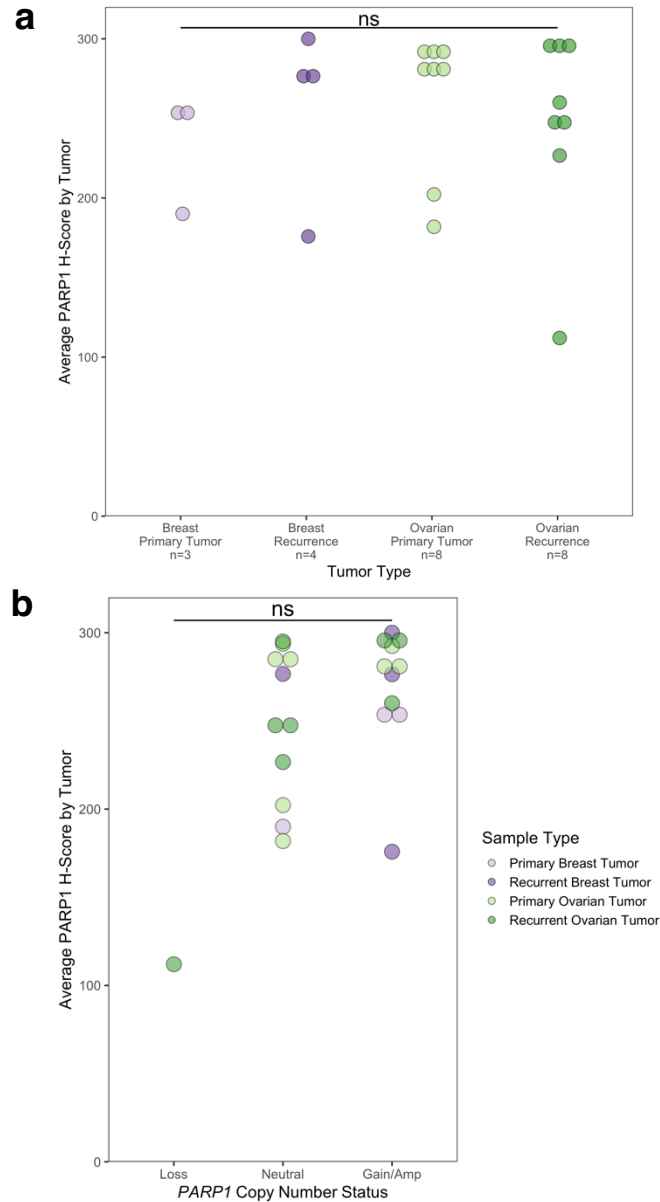


Figure 37: PARP1 protein expression in primary and recurrent tumors. A. PARP1 nuclear positivity (average tumor H-score) by tumor type, for all 23 tumors in tissue microarrays. B. PARP1 nuclear positivity (average tumor H-score) by *PARP1* copy number status (from whole exome sequencing) across all tumors. Groupwise differences were determined by Kruskal-Wallis test, followed by Dunn's test with Bonferroni correction ($\alpha=0.05$).

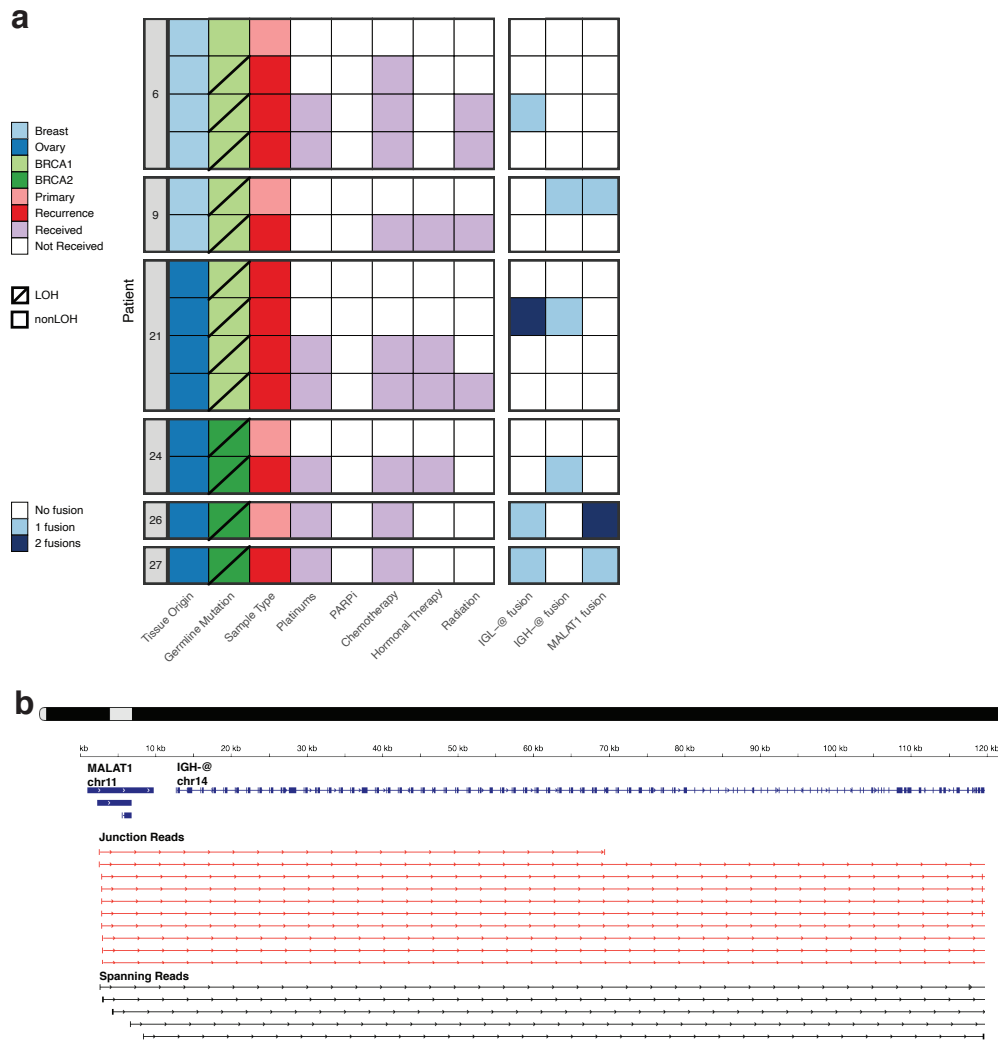


Figure 38: Gene fusions detected by RNA sequencing. A. Clinical characteristics of patients in which *IGH-@*, *IGL-@*, and *MALAT1* fusions were identified. Tumors are displayed in chronological order by patient, with the primary tumor at the top and latest recurrence at the bottom. “1 fusion” refers to a single translocation with one other gene. B. Example of *MALAT1-IGH-@* gene fusion IGV tracks from patient 9. Junction reads (red, middle track) represent split RNA-seq reads used to map the fusion breakpoint. Spanning reads (black, bottom track) represent paired-end reads of fragments that span, but do not directly overlap, the fusion breakpoint.

Discussion

We performed a comprehensive transcriptomic analysis encompassing our uniquely large set of primary and recurrent *BRCA1/2* mutation-associated tumors. Based on clustering analyses, we found that global transcription differed most on the basis of tumor vs. normal status and ER positivity. Tissue origin appeared to delineate samples to a lesser degree than ER status, and *BRCA1* vs. *BRCA2*-derived differences were not observed at all. Surprisingly, tumors from the same patient were not always closely related, underscoring the utility of a groupwise (rather than primary/recurrent pairwise) approach for the RNA-seq arm of this project.

We observed differences in gene expression between primary and recurrent tumors, which varied by tumor type. Ovarian recurrences had increased expression of genes relating to KRAS signaling, epithelial to mesenchymal transition (EMT), immune activation, and angiogenesis, compared to ovarian primary tumors. Conversely, breast primary tumors had higher expression of these gene sets compared to breast recurrences. These differences suggest that the pro-tumorigenic signaling that drives breast primary tumors is shared by ovarian recurrences in *BRCA1/2* mutation carriers. Breast recurrences were only enriched for one gene set over primary tumors, possibly due to high heterogeneity (differences in ER status, treatment exposure, and site) within this group. Of note, our results agree with a prior study which found that primary and recurrent *BRCA1/2* mutation-associated breast tumors differ in gene expression related to $\text{TNF}\alpha$ signaling and EMT⁷¹.

When primary and recurrent tumors were combined by tissue origin, many of the same gene sets were enriched in breast and ovarian tumors overall. These included pathways for *MTOR*, *MYC*, and some inflammatory signaling. Some of the gene sets identified by pathway analysis (JAK/STAT and $\text{TNF}\alpha$ signaling, etc.) could represent

novel therapeutic targets for *BRCA1/2* mutation-associated cancer. We also observed a general loss of lineage-related transcription factors, as evidenced by motif analysis of genes under-expressed in tumors. We propose that loss of cell type-specific transcriptomic programs contributes to primary and recurrent tumor phenotypes, in addition to other pro-tumorigenic signaling pathways observed across *BRCA1/2* mutation-associated tumors. Notably, none of the tumor groups showed differential expression of *CHD4*, *TP53BP1*, or *MAD2L2/REV7* compared to normal tissue. Therefore, we also concluded that *CHD4*, *TP53BP1*, or *MAD2L2/REV7* losses were not a major mechanism of acquired therapeutic resistance in this cohort of tumors.

We also assessed gene fusions as potential drivers of *BRCA1/2* mutation-associated tumors. We did not identify activating fusions involving *ABCB1*, as reported by others^{51,71}. The absence of *ABCB1* hyperactivation is also supported by our finding that all tumor groups had significantly lower *ABCB1* expression than tissue-matched normal specimens, in spite of the *ABCB1* gains noted in Chapter 4. We also found three tumors with fusions involving the lncRNA *MALAT1*. *MALAT1* fusions have been observed in other cohorts of breast and ovarian tumors after PARPi, although none of the tumors bearing *MALAT1* fusions in this cohort received PARPi¹⁴⁹. *MALAT1* has been shown to interface with a variety of pro-tumorigenic signaling pathways, and its expression is correlated with metastasis and poor survival in several tumor contexts¹⁵⁰. Here, we demonstrate that *MALAT1* fusions may drive a small subset of *BRCA1/2* mutation-associated breast and ovarian cancers as well.

Lastly, we identified *PARP1* overexpression as a shared feature of primary and recurrent tumors. We observed >2-fold increases for *PARP1* expression in primary and recurrent breast tumors, as well as ovarian tumor recurrences, relative to normal tissue (all adj. $p < 0.05$). *PARP1* transcript abundance and PARP1 H-score were high across

all tumor types, including those without *PARP1* gains and amplifications. Therefore, we propose that *PARP1* copy number gains represent one explanation for PARP1 overexpression, and that such gains are common in breast and ovarian tumors regardless of *BRCA1/2* status. Our results agree with prior reports of high PARP1 expression in unselected ovarian and receptor-negative breast cancers^{151–153}.

Increased nuclear PARP1 protein has been correlated with decreased relapse-free and overall survival in breast cancer¹⁵⁴. In this context, we hypothesize that increased levels of PARP1 protein may stoichiometrically dilute PARPi, while also potentiating PARP1-dependent alternative DNA repair pathways and transcriptional programs. As PARPi are increasingly used to treat *BRCA1/2* mutation-associated and sporadic breast and ovarian tumors, *PARP1* amplifications and overexpression may represent a biomarker of PARPi response. In this study, we present the first evidence of *PARP1* overexpression as a near-ubiquitous feature of primary and recurrent *BRCA1/2* mutation-associated tumors, a result which may have significant relevance in the clinic.

CHAPTER 6: DIFFERENTIAL TRANSCRIPT USAGE

Introduction

Human cells use alternative splicing to transcribe multiple mRNA transcripts from the same DNA sequence. This phenomenon is estimated to occur in 95% of multi-exon human genes¹⁵⁵. In normal contexts, alternative splicing enables cells to fine-tune, diversify, and expand the proteome encoded by a relatively small number of genes^{156–158}. However, in tumor contexts, alternative splicing is commonly pro-tumorigenic and can contribute to every hallmark of cancer^{157,158}. For example, cancer syndrome genes (tumor suppressors, including *BRCA1/2*) are particularly prone to splice site and exonic mutations that disrupt canonical splicing to abrogate protein expression¹⁵⁹. Such changes in *cis*-acting DNA sequences represent a common route to pro-tumorigenic alternative splicing in cancer. Other well-recognized mechanisms include overexpression and/or aberrant activity of *trans*-acting splicing factors, including RNA binding proteins^{156,157}.

Alternative splicing usually falls into one of several categories: (1) exon skipping, (2) intron retention, (3) mutually exclusive exons, (4) alternative 3' splice sites, or (5) alternative 5' splice sites^{156,158}. Any of these alterations can effectively modulate gene expression via alterations to mRNA (stability, localization) or protein (coding potential, functional domains)^{156,160}. Proteins translated from tumor-specific mRNA transcripts have also been proposed as neoantigens for cancer immunotherapy and/or potential therapeutic targets^{156,158}. To assess the impact of alternative splicing on *BRCA1/2* mutation-associated cancer, we used RNA-seq from primary tumors, recurrent tumors, and normal tissue samples. In particular, we looked for groupwise patterns in alternative splicing, referred to here as differential transcript usage or isoform switching.

Methods

Assessment of differential transcript usage from RNA-seq

Abundance was re-quantified from bam files (see chapter 2 **Methods**) per recommendations in the Stringtie manual for novel isoform detection (<http://ccb.jhu.edu/software/stringtie/index.shtml?t=manual>). Briefly, starting from initial bam files, we ran Stringtie in merged mode (using GENCODE Human Release 19 as reference) to generate a merged gene transfer format (gtf) file for the cohort. We quantified abundance from initial bams using Stringtie in Ballgown mode and the merged gtf file as reference. Differential transcript usage was assessed, and switch plots generated, using IsoformSwitchAnalyzer (v1.10.0). A validation cohort of primary tumors (see below) was analyzed for isoform switching in an independent use of this workflow.

Use of publicly accessible RNA immunoprecipitation sequencing (RIP-seq) datasets

RIP-seq plots were generated using UCSC Genome Browser (hg19 assembly) with the ENCODE RNA Binding Proteins track corresponding to GEO Accession GSE35585. The track pictured for ELAVL1/HuR in **Fig. 41** shows a peak with coordinates chr13:32,973,328-32,973,692, (average enrichment 20.30, FDR $q=4.2 \times 10^{-3}$). The track pictured for PABPC1 in **Fig. 41** shows a peak with coordinates chr13:32,973,450-32,973,693 (average enrichment 31.59, FDR $q=3.7 \times 10^{-3}$).

Acquisition of validation cohort from Penn

For validation of *BRCA2* isoform switching, we evaluated an independent cohort of 42 primary breast and ovarian tumors from patients with germline *BRCA1/2* mutations, also consented under the University of Pennsylvania approved IRB protocol. Eligible patients

met the following criteria: (1) diagnosis of breast or ovarian cancer, (2) available FFPE tumor tissue, and (3) positive genetic test result for pathogenic germline mutation in *BRCA1* or *BRCA2* from a CLIA-approved laboratory. RNA extraction, sequencing, and analyses matched those used for primary/recurrent cohort samples.

Statistical and survival analyses

The adjusted p-value for total *BRCA2* expression was generated by linear modeling with limma (see chapter 5 **Methods**). The adjusted p-values for *BRCA2* isoform usage were computed using DEXSeq within isoformSwitchAnalyzer. Association between sample type and *BRCA2* isoform usage was determined by Mantel-Haenszel chi-squared test with continuity correction ($\alpha=0.05$, 1df).

Overall survival was defined as time from diagnosis to death or last follow-up. Survival metrics were calculated based on manual chart review for each patient. The survival analysis was limited to 67 patients for whom we collected survival data and RNA-seq from at least one tumor (see **Supplementary File 11**). For the 33 ovarian cancer patients, we first used a Cox proportional hazards model to control for patient age at diagnosis, tumor stage at diagnosis, and disease recurrent status as confounding variables. We used the same confounding variables, in addition to ER status, in multivariate Cox proportional hazards analysis of the 34 breast cancer patients. In ovarian cancer patients, the only confounding variable associated with survival was disease recurrence status (a logical variable indicating whether the patient's disease eventually recurred, regardless of our collecting that recurrence). Therefore, we simplified our ovarian Cox proportional hazards model to include patient recurrent status as the only confounding variable. In breast cancer patients, 0/4 confounding variables

tested were significantly associated with survival, so these were discarded in favor of a univariate analysis. Cox proportional hazards analyses were performed and plotted using the Survival (v3.2.7) and Survminer (v0.4.8) R packages. P-values reported on the plots are those corresponding to the isoform expression term itself. P-values for the models were tested for significance by Wald test ($\alpha=0.05$, 1 df for breast, 2 df for ovarian).

Results

Differential *BRCA2* isoform usage in primary and recurrent tumors

We identified an isoform switching event between recurrences and normal tissue, involving two protein-coding *BRCA2* transcripts. Overall, *BRCA2* gene expression was significantly higher in recurrent tumors than normal tissue (adj. $p=7.13 \times 10^{-6}$, **Fig. 39a**). This increase in *BRCA2* expression was found in both breast and ovarian recurrences when compared separately to matched normal tissue (both $\log_{FC} \geq 1.79$ and $\text{adj.} p < 1 \times 10^{-3}$). We found that recurrent tumors preferentially expressed a shorter transcript (ENST00000380152.3, $q=1.7 \times 10^{-4}$), whereas normal samples favored a longer transcript (ENST00000544455.1, $q=6.0 \times 10^{-4}$) (**Fig. 39b**). *BRCA2* ENST00000544455.1 has a longer 3' UTR, and is predicted to be sensitive to nonsense-mediated decay (NMD); whereas with a shorter 3' UTR, *BRCA2* ENST00000380152.3 is predicted to be NMD-insensitive (**Fig. 39c**)¹⁶¹. We observed that expression of each isoform was mutually exclusive in 64 of 66 (97%) RNA-seq samples, with a significant association between sample group (normal, primary tumor, or recurrence) and isoform usage (**Fig. 40**, $p=0.001$). Across all tumor samples, expression of ENST00000380152.3 was slightly more common in breast tumors (10/20; 50%) than in ovarian tumors (14/34; 42%). Tumor type differences were more pronounced within the recurrence group;

ENST00000380152.3 was expressed in 8 of 11 (73%) breast recurrences compared to 12 of 23 (52%) ovarian recurrences. In general, ENST00000380152.3 was expressed more frequently in *BRCA1* mutation-associated tumors (21/42; 50%) than in *BRCA2* mutation-associated tumors (3/12; 25%). This difference was also largest within the recurrence group; ENST00000380152.3 was expressed in 18 of 27 (67%) *BRCA1* mutation-associated recurrences compared to 2 of 7 (29%) *BRCA2* mutation-associated recurrences. There were no apparent associations with treatment history or LOH status.

Of note, the differing region of the *BRCA2* 3' UTR contains binding sites for RNA binding proteins ELAVL1 (HuR) and PABPC1 (both FDR $q < 0.01$, **Fig. 41**), based on RNA immunoprecipitation sequencing (RIP-seq) experiments in human B cells (GEO Accession GSE35585). The presence of HuR binding sites in introns and 3' UTR has been directly associated with increased mRNA stability¹⁶². PABPC1 has been shown to inhibit NMD in mRNA transcripts with long 3' UTRs^{163,164}. Of note, expression of both *PABPC1* and *ELAVL1* genes were significantly increased in this cohort of tumors compared to normal tissue (module 2, **Supplementary File 9**). Our data support the presence of two *BRCA2* transcripts, which may differ in stability and/or localization, as regulated by RNA binding proteins and/or NMD^{165,166}. ENST00000380152.3, the shorter isoform, is preferentially expressed in recurrent tumors (particularly breast) and in association with *BRCA1* mutations.

Validation of *BRCA2* transcript usage and effects on survival

To validate these results, we assessed *BRCA2* isoform expression in an independent cohort of 42 *BRCA1/2* mutation-associated primary breast and ovarian tumors (**Supplementary File 1, Fig. 42**). As in our discovery set, primary tumors in the

validation cohort more commonly expressed the longer isoform (33/42; 79%).

Expression of the shorter isoform ENST00000380152.3 was similarly uncommon in breast vs. ovarian primary tumors (28%, 29%) and *BRCA1* vs. *BRCA2* mutation-associated primary tumors (28%, 31%).

Next, we tested whether expression of the shorter *BRCA2* isoform affected overall survival across 67 *BRCA1/2* mutation carriers (from discovery and validation cohorts) for whom we had survival data and RNA sequenced from at least one tumor (**Supplementary File 11**). We tested the Cox proportional hazards model for significant associations with ER status, age at diagnosis, tumor stage at diagnosis, and patient recurrence status. Out of 34 breast cancer patients assessed, those expressing ENST00000380152.3 in any tumor had significantly worse overall survival compared to patients not expressing this alternative isoform ($p=0.017$, $HR=2.535$, 95% CI 1.179-5.45) (**Fig. 43**). Median overall survival was nearly three years shorter for patients whose tumor(s) expressed ENST00000380152.3 (87 months, 95%CI 71.2 - 121) compared to those whose tumors did not (121 months, 95%CI 100.0 – 159). This effect appeared not to be driven by the association between ENST00000380152.3 expression and disease recurrence, as patient recurrent status was not significantly associated with the Cox proportional hazards model. Further, breast tumors expressing ENST00000380152.3 were evenly split between patients that eventually recurred (54%) and those that did not (46%). Expression of the shorter isoform was not correlated with overall survival in 33 ovarian cancer patients after adjustment for patient recurrence status (**Fig. 44**).

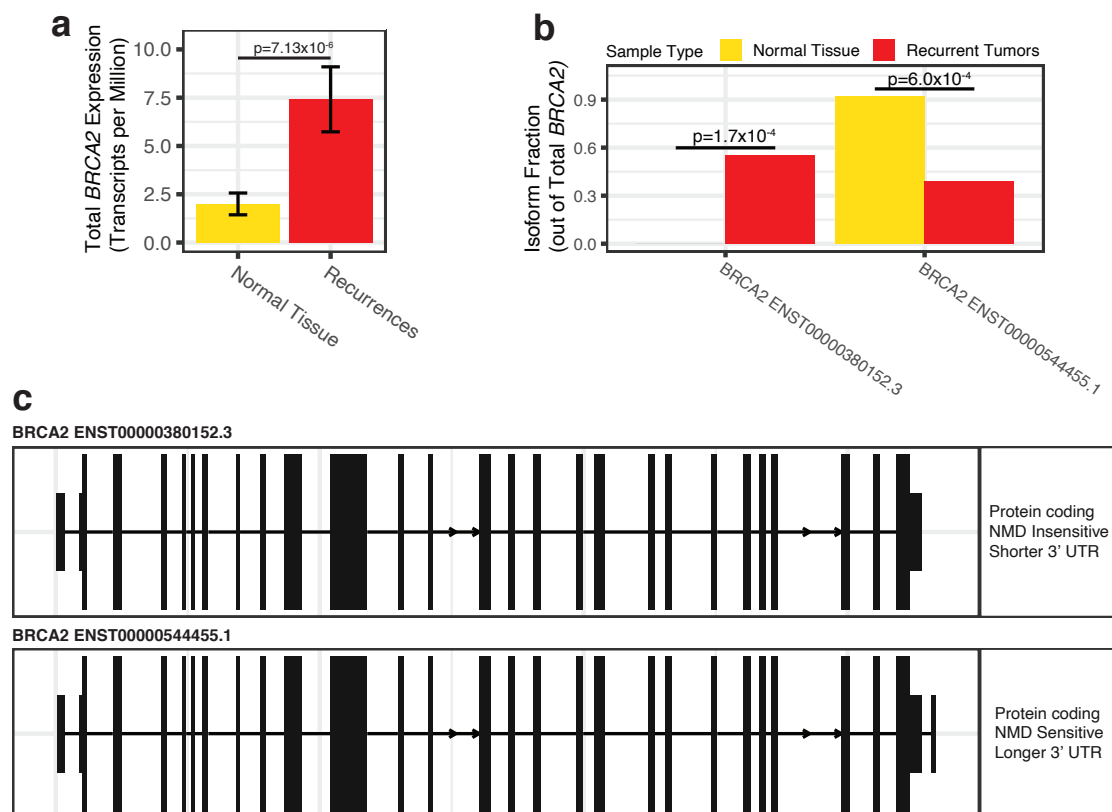


Figure 39: *BRCA2* gene and transcript expression in normal tissue and recurrent tumors. A. Expression of total *BRCA2* (all isoforms) with 95% CI in recurrences vs. normal samples. Adjusted p-value for differential gene expression is based on linear modeling of mean-variance trends (limma). B. *BRCA2* isoform usage (*BRCA2* isoform expression normalized to total *BRCA2* expression) in recurrences vs. normal samples. Isoform switch q values were computed using DEXSeq within isoformSwitchAnalyzer. C. *BRCA2* transcripts involved in isoform switching events. Transcripts are labeled with Ensembl transcript IDs corresponding to human genome release GRCh37.

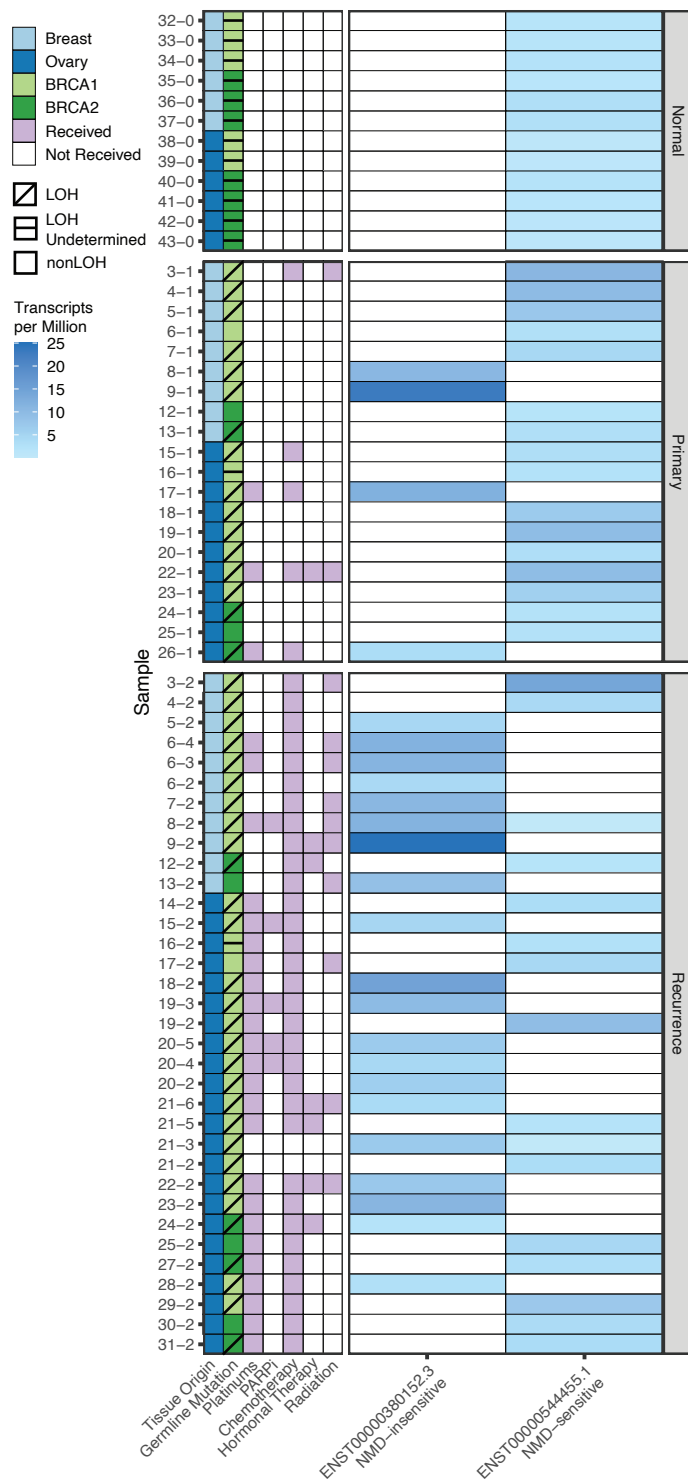


Figure 40: *BRCA2* isoform expression by sample and group for entire RNA sequencing cohort.

Groupwise differences in transcript usage were tested for significance by Mantel-Haenszel chi-squared test with continuity correction ($\alpha=0.05$, 1df). P-value ($p=0.001$) is not displayed. Expression corresponds to transcripts per million.

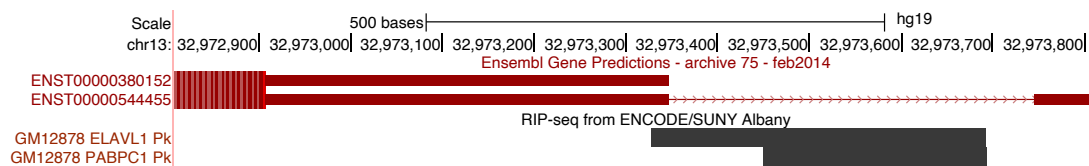


Figure 41: RNA binding proteins in the 3' UTR region of difference between *BRCA2* transcripts. RNA binding protein (ELAVL1/HuR and PABPC1) prediction for *BRCA2* transcripts based on RNA immunoprecipitation sequencing (RIP-seq) in GM12878 (human B cells). View is zoomed in to the region of difference between transcripts (in each transcript's 3' UTR). RIP-seq tracks come from GEO Accession GSE35585; see **Methods**.

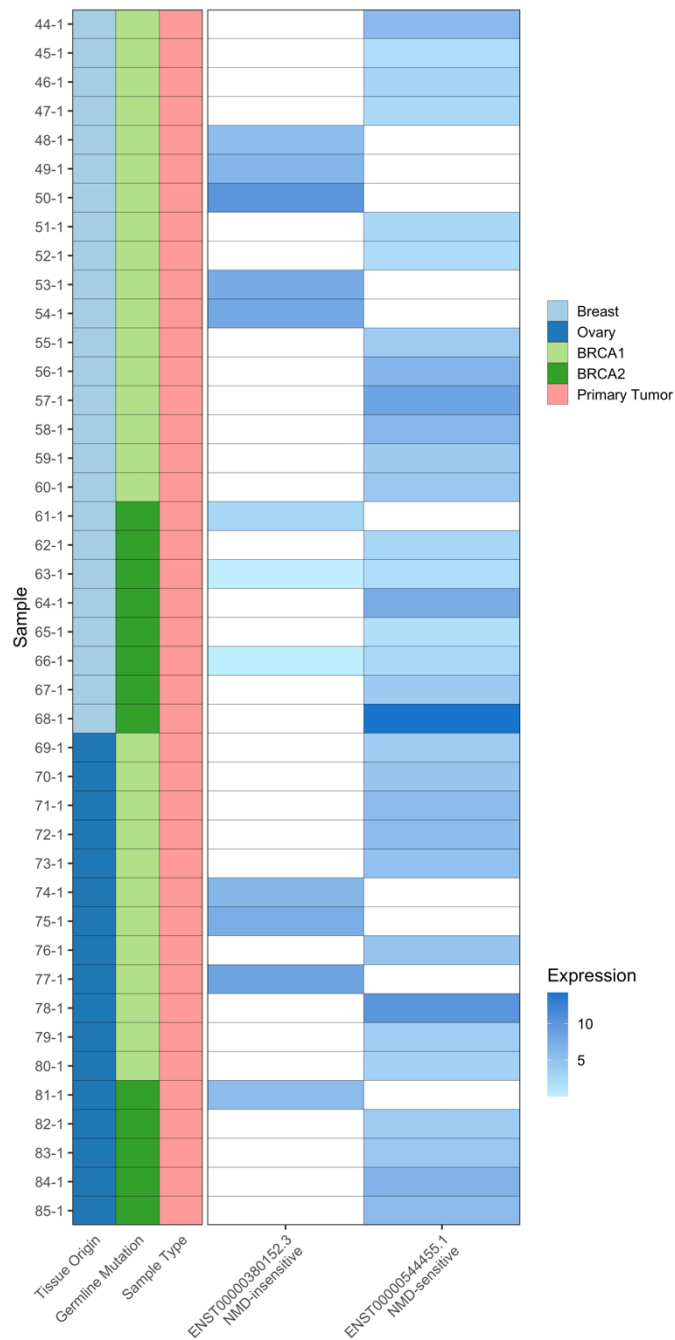


Figure 42: Validation of *BRCA2* isoform detection in an independent cohort. *BRCA2* isoform usage by sample type in a cohort of 42 additional primary breast and ovarian tumors from *BRCA1/2* mutation carriers. None of these tumors were included in the original primary/recurrent cohort for the rest of the study. Expression corresponds to transcripts per million.

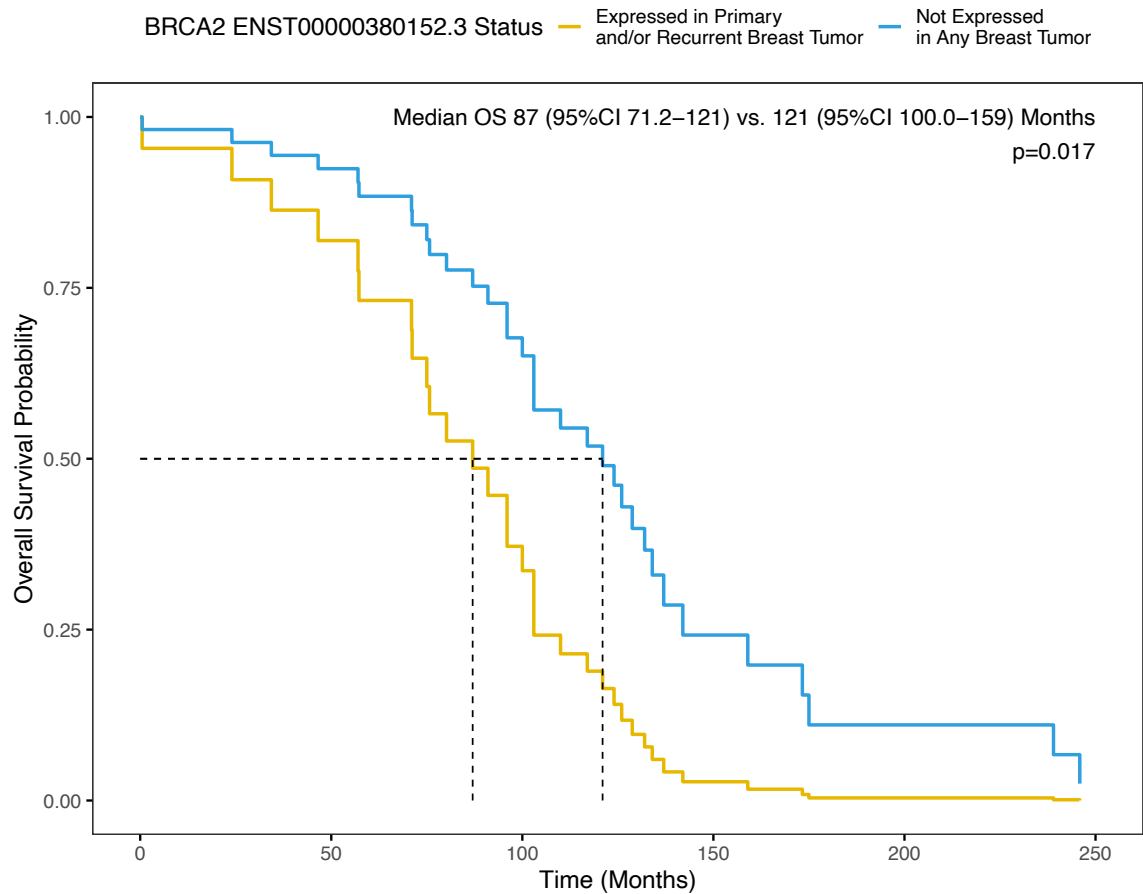


Figure 43: Effect of differential *BRCA2* transcript usage on overall survival in *BRCA1/2* mutation carriers with breast cancer. Survival curve for patients that expressed ENST00000380152.3 in any (primary or recurrent) breast tumor compared to those that did not. Survival proportions and p-value were calculated using a Cox proportional hazards model and tested for significant associations with ER status, age at diagnosis, tumor stage at diagnosis, and patient recurrent status ($\alpha=0.05$, see **Methods**).

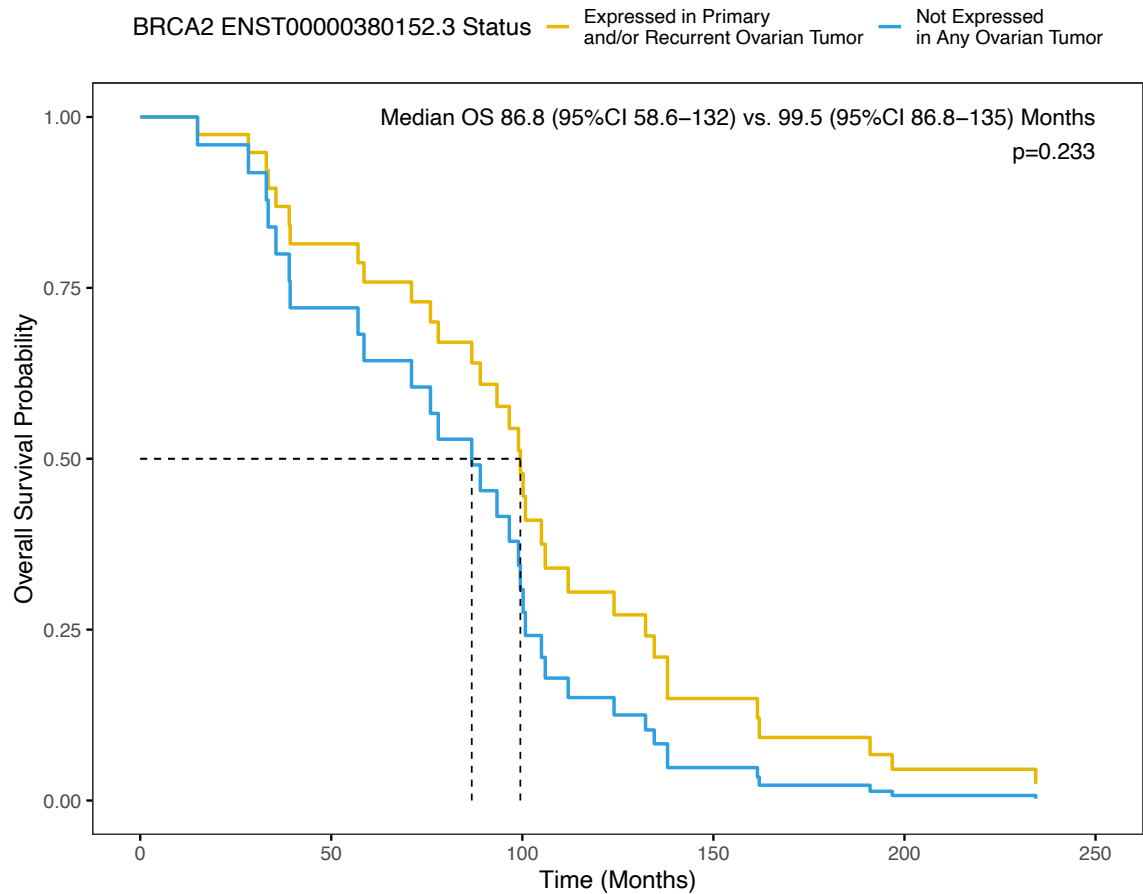


Figure 44: Effect of differential *BRCA2* transcript usage on overall survival in *BRCA1/2* mutation carriers with ovarian cancer. Survival curve for patients that expressed *BRCA2* ENST00000380152.3 in any (primary or recurrent) ovarian tumor compared to those that did not. Survival proportions and p-value were calculated using a Cox proportional hazards model tested for significant associations with ER status, age at diagnosis, and tumor stage at diagnosis; and adjusted for patient recurrent status ($\alpha=0.05$, see **Methods**).

Discussion

We assessed transcriptomic differences in primary and recurrent tumors and found two different isoforms of *BRCA2*. *In silico* predictions suggest that the *BRCA2* transcripts differ in regulation of mRNA stability and/or localization, with the longer form insensitive to NMD^{160,161}. We found that both HuR and PABPC1 (RNA binding proteins) were predicted to bind the implicated region of difference between isoforms in the 3' UTRs; and further, both HuR and PABPC1 were overexpressed in tumors compared to normal tissue. The shorter *BRCA2* isoform was significantly more frequent in recurrent tumors, especially breast tumors, and those from *BRCA1* mutation carriers.

We hypothesize that the isoform switch may represent a mechanism by which *BRCA1* mutation-associated tumors can modulate *BRCA2* transcript stability to retain HR function and enhance tumor survival. Since most pathogenic germline *BRCA2* mutations are truncating mutations, *BRCA2* isoform switching would not likely rescue HR in *BRCA2* mutation-associated tumors with LOH¹⁶⁷. However, in nonLOH *BRCA2* mutation-associated tumors, isoform switching may protect mRNA transcribed from the wild-type *BRCA2* allele. Multiple mechanisms have been demonstrated to restore *BRCA2* function in resistance to platinum and PARPi^{55,56,168–170}. A recent study found that even partial *BRCA2* function, through amplification of truncated *BRCA2*, promotes RAD51 foci formation and PARPi resistance in Capan1 cells¹⁷⁰. The same mutation found in Capan1 cells (*BRCA2* c.5946delT) was associated with two of three *BRCA2* tumors expressing the shorter *BRCA2* isoform¹⁷⁰.

Our survival analysis offers further support for a correlation between *BRCA2* isoform expression and clinical outcomes. We found that breast cancer patients expressing the longer *BRCA2* isoform in any tumor (primary or recurrences) had

significantly improved overall survival (median 87 vs 121 months), suggesting it could be a prognostic biomarker. Taken together, these results suggest that *BRCA1* mutation-associated tumors, as well as *BRCA2* mutation-associated tumors lacking LOH or bearing the *BRCA2* c.5946delT germline mutation, could benefit from enhanced *BRCA2* mRNA stability. We therefore propose that *BRCA2* isoform switching represents a novel tumorigenic driver in *BRCA1/2* mutation-associated cancers.

CHAPTER 7: CASE STUDIES IN THREE PATIENTS

Introduction

We integrated multi-omic results with clinical histories to trace the tumor evolution for three patients of interest. We chose to examine patients with unique genetic features, including LOH transitions, *BRCA2* isoform switching, and a *BRCA1* reversion mutation. Here we present their cases for an in-depth analysis of the relationship between primary and recurrent tumors.

Methods

Curation of multi-omic results and clinical metadata by patient

Clinical metadata (treatment history, outcomes, receptor status) were curated by manual chart review. For somatic variants from whole exome and targeted sequencing (including detection of *BRCA1/2* reversion mutations), see chapter 3 **Methods**. For copy number variation, associated scores, and determination of allele-specific LOH, see chapter 4 **Methods**. For differential *BRCA2* transcript usage, see chapter 6 **Methods**.

Tools used for illustrations

Patient-specific figures were made in BioRender (<https://biorender.com/>). The graphical representation of Patient 20's *BRCA1* reversion mutation was also made in BioRender. Whole exome sequencing tracks for Patient 20 were generated using bam files and the Integrative Genomics Viewer (IGV; IGV_2.8.13 for Mac)⁹⁵. The forward DNA strand is displayed.

Results

Tumor evolution in Patient 20

Patient 20 was a *BRCA1* mutation carrier with an ovarian primary tumor and four recurrences (**Fig. 45**). She had an initial durable response to platinum, then received PARPi (Olaparib) after recurrence #2. Over nine years and multiple lines of therapy, the tumor maintained *BRCA1* LOH, high aneuploidy scores (≥ 13), copy number gains in ABC transporters *ABCB10/11*, *MYC* amplifications ($CN \geq 10$), and *PARP1* gains ($CN=4$) seen in the primary tumor (**Fig. 14a**). However, *BRCA2* isoform usage evolved, with the primary tumor expressing *BRCA2* ENST00000544455.1 (longer isoform) and recurrences expressing ENST00000380152.3 (shorter isoform). Interestingly, recurrence #2 was negative for LOH and copy neutral for *ABCB10/11* (**Fig. 14b**). Based on its divergent genomic features, including distinct LoF mutations in *TP53BP1*, *NF1*, *TSC1*, and *ATM*, recurrence #2 appears to have evolved from a distinct clone.

Following 12 months of PARPi treatment, a *BRCA1* reversion is readily detectable in recurrences #3 and #4. The reversion was a 15bp in-frame deletion encompassing the site of the germline nonsense mutation (**Fig. 46, 47**). The reversion is not present in earlier specimens from this patient (primary tumor or post-platinum recurrences #1 or #2) but appears to have developed after PARPi exposure. In line with the *BRCA1* restoration expected to result from this reversion, the patient eventually progressed on PARPi^{54,60,171}.

Tumor evolution in Patient 13

Patient 13 was a *BRCA2* mutation carrier with TNBC, whose primary tumor had LOH but yielded a nonLOH recurrence (**Fig. 48**). This observation suggests that the recurrence was populated by outgrowth of nonLOH sub-clones. We interrogated her case, since she was the only patient who underwent LOH reversal without a concomitant

decrease in HRD score (**Fig. 15**, left panel). Patient 13's recurrence had increased TMB, aneuploidy, and HRD scores, indicating accumulation of genomic instability and mutations over the tumor's lifetime. This observation contrasts the trends observed for the cohort as a whole, for which HRD scores, aneuploidy, and TMB were consistent from primary tumor to recurrence (**Fig. 8**, **Fig. 16**).

The tumor also showed *BRCA2* isoform switching and development of a *MYC* amplification, as have been widely observed in prior studies of *BRCA1/2* mutation-associated cancers^{124–126}. Interestingly, the recurrence had copy number losses and mutations in HR and other DNA damage response genes, including *MUTYH*, *RAD54L*, *RAD52*, and *RAD50*. *RAD52* specifically is required for survival of *BRCA2*-deficient cells¹⁷². Accordingly, this result suggests that *RAD52* proficiency was maintained in the setting of biallelic *BRCA2* loss (primary tumor with LOH) but dispensable after LOH reversal (nonLOH recurrence).

Tumor evolution in Patient 6

For Patient 6, a *BRCA1* mutation carrier, we evaluated one primary tumor and three recurrences of TNBC (**Fig. 49**). We chose to interrogate her case because she was one of three patients whose tumors developed LOH upon recurrence. In addition to allele-specific LOH, her recurrences also demonstrated a *BRCA2* isoform switch (**Fig. 13b**). Two recurrences had copy number gains in *ABCB10/11*.

Over the course of tumor progression, recurrences underwent loss of DNA damage response and repair genes, including *ATR*, *ATM*, *RAD50*, *MRE11*, *RAD18*, *ERCC1/2/6*, *XRCC1*, *MSH2/6*, and *XPC*. Disruption of these genes (via LoF mutations or copy number losses) could abrogate several mechanisms of DNA repair. The

deleterious alterations were not conserved over the course of this patient's tumor evolution. Instead, they clustered into shared pathways, indicating a convergence of many different mechanisms to facilitate defective DNA damage response and repair in each recurrence.

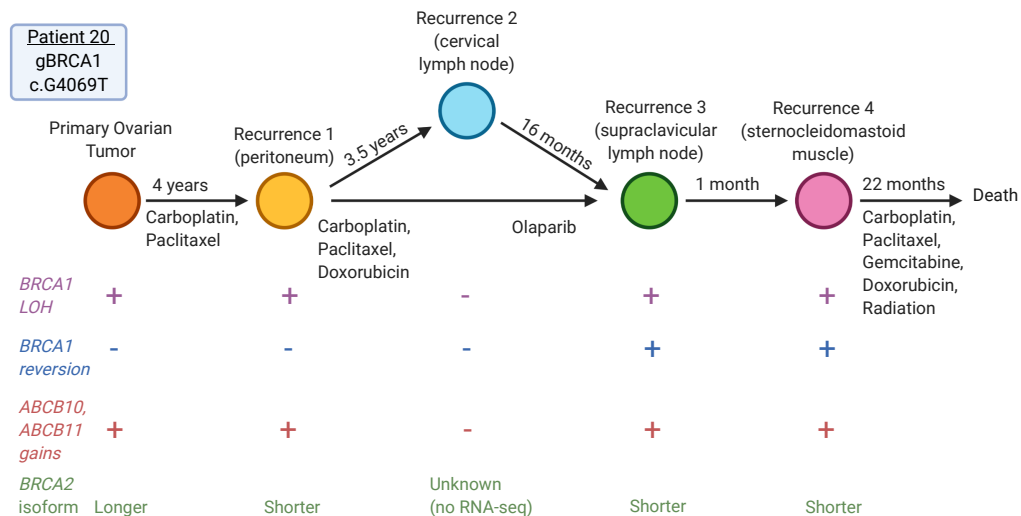


Figure 45: Genetic and clinical features in Patient 20, a *BRCA1* mutation carrier with ovarian cancer. Display is limited to tumors sequenced for this study. Copy number gains refer to genes with total copy number ≥ 4 . “Longer” *BRCA2* refers to ENST00000544455.1; “Shorter” *BRCA2* refers to ENST00000380152.3.

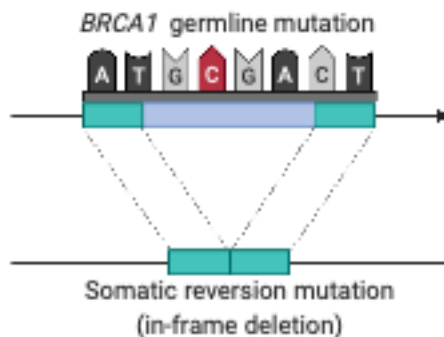


Figure 46: Somatic reversion of *BRCA1* in Patient 20. Graphical representation of in-frame somatic deletion of germline nonsense mutation in *BRCA1* (mutation pictured is not Patient 20’s actual germline mutation).



Figure 47: Sequencing tracks demonstrating *BRCA1* reversion after PARPi in Patient 20. IGV tracks from WES of germline (top track) and tumor DNA from Patient 20. Tumor tracks are displayed in chronological order from top to bottom. For each sample, forward strand alignment and coverage tracks are displayed around the site of the patient's germline *BRCA1* c.G4069T (p.E1357X) mutation. G>T mutation appears as C>A because *BRCA1* lies on reverse strand. See **Figure 45** for treatment exposure by tumor in Patient 20.

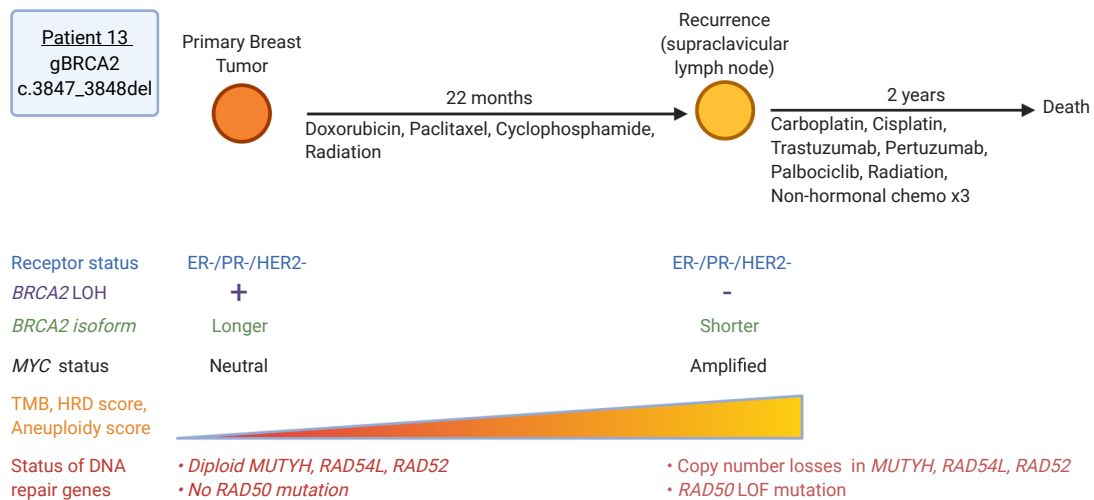


Figure 48: Genetic and clinical features in Patient 13, a *BRCA2* mutation carrier with breast cancer. Display is limited to tumors sequenced for this study. Copy number losses refer to genes with total copy number of 0 or 1; copy number amplifications refer to genes with total copy number ≥ 6 . Loss of function (LoF) mutations reported here are from high-depth targeted sequencing. “Longer” *BRCA2* refers to ENST00000544455.1; “Shorter” *BRCA2* refers to ENST00000380152.3.

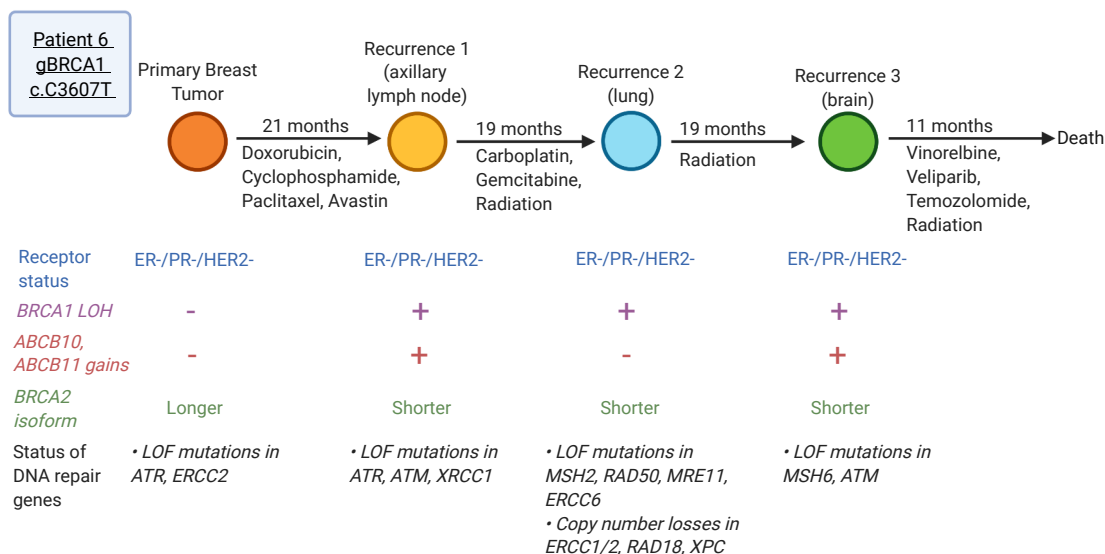


Figure 49: Genetic and clinical features in Patient 6, a *BRCA1* mutation carrier with breast cancer. Display is limited to tumors sequenced for this study. Copy number losses refer to genes with total copy number of 0 or 1; copy number gains refer to genes with total copy number ≥ 4 . Loss of function (LoF) mutations reported here are from high-depth targeted sequencing. “Longer” *BRCA2* refers to ENST00000544455.1; “Shorter” *BRCA2* refers to ENST00000380152.3.

Discussion

In this section, we examined tumor evolution in three patients from the primary/recurrent cohort. We focused specifically on patients with LOH transitions (LOH to nonLOH in Patients 20 and 13, nonLOH to LOH in Patient 6). We also chose cases that illustrated *BRCA2* isoform switching from primary tumor to recurrence, among other unique genetic features.

We presented Patient 20’s case because her tumors demonstrated the only

BRCA1 reversion event observed in the cohort. While *BRCA1/2* reversions are a widely described mechanism to platinum and PARPi resistance, they tend to be reported in case studies⁶⁰. A recent meta-analysis of these studies noted the low prevalence of *BRCA1/2* reversions in larger cohorts free from selection bias⁶⁰. This meta-analysis also found that *BRCA1/2* reversion mutations were most common in ovarian cancer patients (like Patient 20), likely due to routine platinum and PARPi use for this cancer type⁶⁰. Reversions primarily take the form of small, single-exon deletions near the site of a truncating germline mutation, as was also observed in Patient 20⁶⁰. Interestingly, Patient 20 did not develop her *BRCA1* reversion after initial treatment with platinum alone, instead only demonstrating reversion after PARPi. Her case illustrates that platinum alone may not facilitate *BRCA1* reversion (as detectable by bulk sequencing).

Another interesting feature of Patient 20's case was the retention of several genetic features, despite extensive therapy exposure and time. While one tumor (recurrence #2) appeared to follow its own branch of tumor evolution, other recurrences demonstrated consistent *ABCB10/11* gains and use of the shorter *BRCA2* transcript. Similarly, we noticed that Patient 6's multiple recurrences retained their triple-negative receptor status, *BRCA1* LOH, and *BRCA2* transcript usage over the course of several years and treatments (including platinum). We presented these cases to illustrate that multi-omic features may be retained in multiple recurrences from the same patient.

A key conserved feature in Patient 6's case was the disruption of DNA repair pathways. We found LoF mutations and copy number losses predicted to abrogate translesion synthesis, nucleotide excision repair, interstrand crosslink repair, transcription-coupled repair, single-strand break repair, mismatch repair, and double-strand break repair^{173–176}. Mutations in these genes have been reported in gliomas and a

variety of carcinomas, including lung, gastric, bladder, colorectal, and prostate^{177–183}. Loss of RAD50, ATM, ATR, and XPC in particular could mediate dysregulation of the DNA damage response and associated apoptosis^{174–176}. While the genetic mechanisms of these losses were not conserved across tumors, the pathways they affected were consistent. Taken together, our results again indicate that primary and recurrent tumors share many of the same genetic features and drivers, even after multimodal therapy and the passage of time. This finding suggests that a personalized medicine approach could be useful for *BRCA1/2* mutation carriers with cancer, whose primary tumors (and earlier recurrences) could be collected and sequenced to predict subsequent responses to therapy.

Lastly, we chose to present Patient 13's case because they were an outlier in the subset of patients with LOH transitions. Unlike other patients with LOH reversal, Patient 13's nonLOH recurrence had a higher HRD score than the LOH primary breast tumor. After examining the case further, we found that aneuploidy scores and TMB were also higher in the patient's recurrence, suggesting that mounting genomic abnormalities may have contributed to her recurrence. We identified copy number losses and LoF mutations in genes required for HR and other DNA repair mechanisms, which represent a potential explanation for increased mutations and chromosomal scarring. RAD52 specifically was lost only after the LOH reversal, in line with the observation that dual RAD52 and BRCA2 deficiency is synthetic lethal¹⁷². We presented Patient 13's case because it illustrates an alternative route to recurrence, mediated by loss of key DNA repair proteins and associated DNA lesions.

CHAPTER 8: CONCLUSIONS AND FUTURE DIRECTIONS

Conclusions

The conclusions for this study largely fall into two categories: multi-omic features that evolved from each primary tumor to recurrence, and features that remained consistent during the entire course of disease.

Conserved multi-omic features

Given the heavy pre-treatment and passing of time, we were surprised by the number of features that did not change in paired tumors. For most (20/27) patients, LOH status was concordant between primary and recurrent tumors, indicating that frontline treatments (including platinum) do not appear to select for nonLOH clones in most cases. While LOH reversals can occur, they were not a major mechanism of platinum or PARPi resistance in this cohort. Markers of genomic abnormalities (HRD and aneuploidy scores, tumor mutational burden) were also consistent between tumors, indicating that the global accumulation of DNA damage does not drive recurrence for most. We also found that primary and recurrent tumors shared many likely drivers (*TP53* and *AKT1* mutations, *MYC* amplifications, *HDAC* deletions). *PARP1* copy number gains, *PARP1* mRNA overexpression, and PARP1 protein overexpression were also major shared features across all tumors in the cohort.

Primary and recurrent tumors also shared many global 'omic patterns. First, each group accumulated deleterious mutations in many of the same pathways, pointing to potentially shared dysregulation of DNA repair processes, Wnt β -catenin signaling, and mitosis. At the level of an individual patient, we also observed that Patient 6's primary tumor and multiple recurrences underwent somatic gene losses affecting several of the

same DNA damage response pathways. Perhaps the strongest evidence for primary/recurrent similarities came from RNA-seq (see Chapter 5). Even though we did a groupwise analysis (rather than assessing tumor pairs by patient), we found a surprising degree of concordant transcription between primary and recurrent tumors. In fact, expression patterns were so similar that linear modeling did not identify any “significantly altered” genes between these groups. Further, gene-level counts per million were highly correlated between primary and recurrent tumors stratified by tumor type, as evidenced by their linear relationship ($R^2 > 0.96$ for both ovarian and breast).

These observations have potential clinical relevance for recurrent disease in *BRCA1/2* mutation carriers. Specifically, they suggest that primary tumors could be utilized to predict therapeutic vulnerabilities in recurrences, since many multi-omic tumor drivers are shared. As an example, *PARP1* gains were typically conserved (or developed into amplifications) during the process of recurrence on a patient-by-patient basis. If *PARP1* overexpression informs tumor phenotypes (as discussed in Future Directions, below), then the *PARP1* copy number status of a patient’s primary tumor could represent a key biomarker of PARPi response in late-stage disease. In theory, primary tumors could also be used to predict well-characterized biomarkers of therapeutic response (HRD score, tumor mutational burden) for the entire course of a patient’s cancer.

Evolution of multi-omic features

A major goal of this study was to identify multi-omic features that distinguish recurrent *BRCA1/2* mutation-associated tumors from matched primary tumors. One interesting observation from RNA-seq was that matched primary/recurrent tumors were

not always most closely related; instead, ER status may play the largest role in delineating tumor transcriptomic programs. Across the cohort, primary and recurrent *BRCA1/2* mutation-associated tumors had many differences at the level of individual genes and pathways. Primary tumors showed copy number gains affecting a higher number of pathways, including those for UV response, innate immune signaling, and angiogenesis. Conversely, recurrences showed distinct copy number gains affecting MTORC1 signaling. Recurrences had significant copy number amplifications in ABC transporter genes *ABCB10* and *ABCB11*, which may contribute to broad therapeutic resistance in this setting¹²². LoF mutations in primary and recurrent tumors clustered into several of the same pathways, but several additional pathways exclusively accumulated deleterious variants in recurrences. Results from both LoF mutations and copy number losses pointed to dysregulated Hedgehog signaling that was also specific to recurrences. Pathway analysis of RNA-seq also indicated that primary and recurrent tumors differed in gene expression from these pathways, with the direction of variation dictated by tumor origin. Taken together, our multi-omic results point to differences in tumor metabolism, as well as cellular identity, growth programs, and survival. They further suggest that different cell signaling programs are beneficial at different stages of tumorigenesis, which could inform use of targeted therapies in the future.

As mentioned above, most primary tumors maintained their LOH status upon recurrence, including in multiple recurrences. However, seven patients had discordant LOH status within their set of tumors. These LOH transitions were corroborated by concomitant changes in HRD score and variant allele fraction of germline SNPs (see Chapter 4). In three *BRCA1/2* mutation carriers, breast tumors demonstrated *BRCA1/2* LOH for the first time in recurrences. These results suggest that LOH is not always an

early event in *BRCA1/2* mutation-associated tumors but may contribute to tumorigenesis at later stages as well. We also found the opposite effect in four patients, for whom primary tumors with LOH yielded nonLOH recurrences.

Prior studies have noted LOH reversals after platinum treatment in *BRCA1* mutation-associated ovarian tumors, which we observed here in two patients. We also observed apparent LOH reversals in *BRCA2* mutation carriers with breast cancer, following chemotherapy and radiation. Three of four patients with LOH reversals had lower HRD scores, in contrast to the conventional wisdom that genomic scarring can only increase with tumor progression. Our findings suggest that LOH reversals may constitute a mechanism by which breast and ovarian tumors restore *BRCA1/2* proficiency, similar to the effect of *BRCA1/2* reversion mutations. While *BRCA1/2* reversions are the best-characterized mechanism of acquired therapeutic resistance in *BRCA1/2* mutation-associated tumors, we found that LOH reversal (four patients) was a more common route to *BRCA1/2* restoration than reversion mutations (one patient) in this cohort.

Lastly, primary and recurrent tumors differed by *BRCA2* transcript usage. Normal tissue from *BRCA1/2* mutation carriers expressed a longer *BRCA2* transcript, which was also favored by primary tumors. Conversely, recurrences expressed a shorter *BRCA2* transcript. Since each *BRCA2* isoform is protein-coding, and the isoforms differ in 3' UTR length, we hypothesize that recurrent tumors express the shorter isoform as way to modulate *BRCA2* mRNA stability or localization. Based on our survival analyses, this mRNA modulation appears to be pro-tumorigenic in breast tumors. While functional studies remain to be done (see Future Directions), we posit that *BRCA2* isoform

switching may represent a novel driver of recurrence in *BRCA1/2* mutation-associated cancers.

Limitations

Historically, studies of *BRCA1/2* mutation-associated tumor pairs have been under-powered. This is largely due to difficulties collecting matched primary and recurrent tumors, which are often banked by different hospitals and collected years apart. Although this cohort represents the largest of paired primary and recurrent *BRCA1/2* mutation-associated tumors published to date, the number of patients was still relatively small (27 patients). As a result, we were limited in our ability to identify correlations between genomic or transcriptomic features and treatment or pathogenic variant type. Treatment was especially heterogeneous for the patients with breast cancer, which is typically more varied than ovarian cancer (due to detection at different stages, receptor status, etc.). Thus, we focused our efforts on identifying similarities and differences between paired tumors, across the cohort and regardless of treatment history. Our rationale was that this approach might identify more robust 'omic features that define unique biology in *BRCA1/2* mutation-associated tumors, while maximizing statistical power. Validation of key results in an independent cohort will be critical, but at the moment, no such cohort is publicly accessible.

In addition to sample size, our approach could not address tumor heterogeneity. Many recurrent tumor specimens are small (lymph nodes, cell blocks from ascites, etc.). Conversely, many primary tumor FFPE blocks have scant tumor remaining, since they have been sectioned heavily for diagnostic and/or research purposes. To maximize the number of tumor pairs while minimizing cost, we only sequenced DNA or RNA from one block per sample. A better approach would have been to sequence samples taken from

multiple sites (each of several FFPE blocks from the same surgery) where possible. While we attempted to use the same DNA samples for whole exome and targeted sequencing, and RNA from the same FFPE tumor block for RNA sequencing, it was not always possible due to sample availability. Lastly, we employed bulk sequencing techniques, so tumor heterogeneity may have obscured meaningful trends when comparing multi-omic results from the same tumor. In particular, bulk sequencing from the primary tumor may not have captured the clone(s) underlying recurrent disease.

This study would have benefitted from an additional analysis of *BRCA1* promoter methylation, which was not possible due to scarcity of remaining tumor DNA. This limitation is an important caveat for the LOH results, since *BRCA1* promoter methylation represents another possible route to biallelic *BRCA1* loss in *BRCA1* mutation-associated tumors. To address the potential for *BRCA1* LOH by promoter methylation, we used RNA-seq to confirm *BRCA1* expression in all nonLOH *BRCA1* mutation-associated tumors. Additionally, *BRCA1* promoter methylation does not always abrogate *BRCA1* expression^{184–186}. These points aside, an ideal study design would have included assessment of *BRCA1* promoter methylation status for all *BRCA1* mutation-associated tumors.

Future Directions (Bench Experiments)

Our key results would benefit from experimental validation. In particular, future directions will include functional characterization of *PARP1* gains and amplifications. *PARP1* overexpression has been observed in ovarian and receptor-negative breast cancers, and associated with decreased survival in breast cancer^{151–154}. However, to our knowledge, no one has assessed tumor phenotypes associated with this overexpression. To address this gap in knowledge, our next steps will include inducing

high PARP1 protein expression (via viral or plasmid-based vector) and assessing the effects on breast and ovarian cancer cells. Specifically, we will test whether PARP1-high cells differ with regards to growth, survival, and resistance to PARPi or other genotoxic therapies. We will perform these experiments in both *BRCA1/2* mutation-associated and *BRCA1/2* wild-type cell lines, to determine whether any phenotypes are informed by *BRCA1/2* deficiency.

Another important next step is to validate the *BRCA2* isoform switching event by orthogonal methods. To do so, we are currently performing isoform-specific qRT-PCR on remaining tumor and normal RNA samples. Our qRT-PCR primers are designed to detect *GAPDH* (endogenous control), the *BRCA2* gene body (positive control; should be detected regardless of *BRCA2* isoform), and the end of the 3' UTR in the longer transcript (should be detected specifically in samples expressing this isoform). Additionally, we could re-section remaining FFPE tumor blocks from the cohort and use an RNAscope to visualize expression and localization of predicted *BRCA2* mRNA transcripts.

Although we were able to determine that the *BRCA2* isoform switching event was associated with decreased survival in patients from two sample sets, we will also perform mechanistic studies to confirm that 3' UTR differences modulate *BRCA2* mRNA. We will create two transgenes containing an inducible promoter, *GFP* gene body, and each of the *BRCA2* isoforms' 3' UTRs. We will express each transgene in breast or ovarian cancer cells, induce transcription, then inhibit transcription globally with Actinomycin D. Following this transcription block, we will use immunofluorescent microscopy to assess GFP protein expression as a readout of each transgene's mRNA half-life. In this way, we can determine the effects of each *BRCA2* 3' UTR on mRNA

stability. We will also use the RNAscope to assess 3' UTR-dependent effects on mRNA localization *in vitro*.

In a preliminary analysis of BRCA2 function after *BRCA2* isoform switching, we used multiplex immunofluorescent microscopy, specifically the co-detection by indexing (CODEX) platform¹⁸⁷. Our lab has stained tissue microarrays containing primary and recurrent tumors (used for PARP1 IHC, see chapter 5) with a 40-plex panel of antibodies optimized for breast and ovarian cancer. This panel detects RAD51 foci, which serve as a readout of BRCA2 function (see Chapter 1) since BRCA2 itself is poorly detected by existing antibodies. In a preliminary analysis of 20 tumors, we found that the percent of RAD51+ cells was slightly higher in tumors expressing the shorter *BRCA2* isoform (ENST00000380152.3) compared to those with the longer isoform (ENST00000544455.1; data not shown). While not statistically significant, this result suggests that *BRCA2* isoform switching yields functional BRCA2 protein, and potentially higher BRCA2 expression.

Based on conversations with GenScript and Dr. Greenberg's lab, we concluded that *BRCA2* isoform switching would be infeasible to model *in vitro* with CRISPR. Technical difficulties are mainly due to the large size of *BRCA2*, as well as the high degree of similarity between the isoforms of interest. Instead, additional functional validation experiments will test the effects of BRCA2 overexpression, which may be one result of *BRCA2* isoform switching. Using BRCA1-deficient breast cancer cell lines, we will assess whether BRCA2 overexpression increases cell growth; survival; or resistance to platinum, PARPi, and other genotoxic therapies. We will prioritize these experiments in breast cancer cell lines (rather than ovarian), since our survival analyses suggest that *BRCA2* isoform switching mediates a more pro-tumorigenic phenotype in this context.

Lastly, we will assess differences in immune milieu between primary and recurrent tumors. The CODEX panel of antibodies contains markers for various immune cell populations and immune checkpoint proteins. We will use CODEX results to characterize populations of innate immune cells (macrophages, dendritic cells, etc.) and T cells (CD8+, CD4+, etc.), since these can all contribute to cancer progression. Prior studies have not assessed expression of immune checkpoint proteins (including PD-1, PD-L1, and CTLA-4) over the lifespan of *BRCA1/2* mutation-associated tumors, and we aim to fill this gap as well. This area of investigation has immediate clinical relevance, since several clinical trials currently aim to assess efficacy of immune checkpoint blockade on *BRCA1/2* mutation-associated breast and ovarian cancers^{188–190}. Previous trials have evaluated immune checkpoint inhibitors as single agent therapies and in combination with PARPi for recurrent *BRCA1/2* mutation-associated tumors. However, objective response rates vary for each therapeutic approach in *BRCA1/2* mutation carriers with breast or ovarian cancer^{191–195}. Therefore, there is a pressing need to characterize immune cell infiltrate and expression of checkpoint proteins in *BRCA1/2* mutation-associated tumors and identify markers associated with response to immunomodulatory therapies.

Future Directions (Bioinformatics)

First and foremost, we must validate key results (LOH transitions, *PARP1* gains, *BRCA2* isoform switching) in an independent cohort of paired primary and recurrent *BRCA1/2* mutation-associated breast and ovarian tumors. While this is not possible due to a lack of appropriately sized and multi-omic datasets, we are hopeful that appropriate validation cohorts will be made publicly accessible in the near future.

We could also extend this study through a number of additional transcriptomic

analyses. One option is to do a paired RNA-seq analysis, as a complement to the groupwise analyses presented here. We chose to focus on groupwise transcriptomic comparisons for the purposes of this study, both to maximize power and to explore tumor/normal comparisons. However, existing RNA-seq methods (EdgeR, etc.) could be used to compare gene expression within primary/recurrent tumor pairs from 19 patients in this cohort. This additional analysis could identify novel trends in gene expression that are specific to recurrences at the level of an individual patient.

Other RNA sequencing analyses could be used to confirm and extend our *BRCA2* isoform switching results. First, we will assess whether differential *BRCA2* transcript usage is detected by another predictive algorithm, MAJIQ¹⁹⁶. We will also determine whether *BRCA2* isoform switching is part of a global alternative splicing program, since statistically significant isoform switches were also identified in other biologically relevant genes (*CDKN2A*, *CHEK2*, *DNMT1*, *DNMT3A*). To do so, we will use PEGASAS, a program that integrates differential transcript usage and gene expression data to determine whether large-scale alternative splicing programs are driven by oncogenic signaling (MYC, MTOR)¹⁹⁷. We found that both MYC and MTOR signaling were elevated in primary and recurrent tumors relative to normal tissue (**Fig. 32**), so this analysis could yield a potential mechanism for the *BRCA2* isoform switching observed.

Lastly, we will use tumor and germline WES for neoantigen prediction. We have already used xHLA to determine HLA type from germline WES for all 27 patients in the cohort¹⁹⁸. Next, we will use antigen.garnish to integrate patient HLA type with somatic mutations to identify high-quality neoantigens for each tumor¹⁹⁹. This analysis will enable us to check for novel, shared neoantigens across the cohort. It will also grant further insights into response to checkpoint blockade in *BRCA1/2* mutation carriers.

Data and Code Availability

Tumor WES and RNA-seq data, as well as RNA-seq from normal tissue samples, will be deposited in NCBI Sequence Read Archive (SRA, <https://www.ncbi.nlm.nih.gov/sra>).

All supplementary data files referenced in the text are described in the Appendix and publicly accessible at

https://figshare.com/articles/thesis/Supplementary_Data_Files/14637513/3.

All custom code associated with this thesis is available at

<https://github.com/nathanson-lab/BRCA-PrimaryRecurrent>.

All custom code associated with HRD and aneuploidy score calculation is available at

<https://github.com/maxwell-lab/HRDex>.

APPENDIX

Supplementary File Guide

All supplementary data files (below) are publicly accessible at
https://figshare.com/articles/thesis/Supplementary_Data_Files/14637513/3

File name: Supplementary_File1.xlsx

Description: Clinical metadata and sequencing performed by tumor (Basser cohorts)

File name: Supplementary_File2.xlsx

Description: Tumor mutational burden, HRD and aneuploidy scores by tumor

File name: Supplementary_File3.xlsx

Description: Gene set enrichment analysis from whole exome sequencing (mutations and copy number variation)

File name: Supplementary_File4.xlsx

Description: MutSigCV results from primary and recurrent tumor cohorts

File name: Supplementary_File5.xlsx

Description: Results from all GISTIC analyses

File name: Supplementary_File6.xlsx

Description: *PARP1* copy number in primary/recurrent and TCGA cohorts

File name: Supplementary_File7.xlsx

Description: TCGA tumors by group

File name: Supplementary_File8.xlsx

Description: Gene set enrichment analysis from RNA sequencing

File name: Supplementary_File9.xlsx

Description: Gene modules identified by RNA sequencing; Gene Ontology results and Transcription Factor Binding Site Motifs identified in gene modules

File name: Supplementary_File10.xlsx

Description: Gene fusions (FusionInspector output) involving *MALAT1* and immunoglobulin genes

File name: Supplementary_File11.xlsx

Description: *BRCA2* isoforms expression by tumor and clinical metadata used for survival analyses

BIBLIOGRAPHY

1. Roy, R., Chun, J. & Powell, S. N. BRCA1 and BRCA2: Different Roles in a Common Pathway of Genome Protection. *Nat. Rev. Cancer* **12**, 68–78 (2012).
2. Zhao, W., Wiese, C., Kwon, Y., Hromas, R. & Sung, P. The BRCA Tumor Suppressor Network in Chromosome Damage Repair by Homologous Recombination. *Annu. Rev. Biochem.* **88**, annurev-biochem-013118-111058 (2019).
3. Prakash, R., Zhang, Y., Feng, W. & Jasin, M. Homologous Recombination and Human Health: The Roles of BRCA1, BRCA2, and Associated Proteins. *Cold Spring Harb. Perspect. Biol.* **7**, (2015).
4. Lord, C. J. & Ashworth, A. BRCAness revisited. *Nat. Rev. Cancer* **16**, 110–120 (2016).
5. Hall, J. M., Lee, M. K., Newman, B., Morrow, J. E., Anderson, L. A., Huey, B. & King, M.-C. Linkage of Early-Onset Familial Breast Cancer to Chromosome 17q21. *Science* (80-.). **250**, 1684–1689 (1990).
6. Wooster, R., Neuhausen, S. L., Mangion, J., Quirk, Y., Ford, D., Collins, N., Nguyen, K., Seal, S., Tran, T., Averill, D., Marshall, G., Narod, S., Lenoir, G. M., Lynch, H., Feunteun, J., Cornelisse, C. J., Menko, F. H., Daly, P. A., Ormiston, W., *et al.* Localization of a Breast Cancer Susceptibility Gene , BRCA2 , to Chromosome 13q12-13. **265**, 2088–2090 (1994).
7. Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., Micklem, G., Barfoot, R., Hamoudi, R., Patel, S., Rice, C., Biggs, P., Hashim, Y., Smith, A., Connor, F., Arason, A., *et al.* Identification of the breast cancer susceptibility gene BRCA2. *Nature* **378**, 789–792 (1995).
8. Miki, Y., Swensen, J., Shattuck-eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L. M., Ding, W., Bell, R., Rosenthal, J., Hussey, C., Tran, T., McClure, M., Frye, C., Hattier, T., Phelps, R., Haugenstrano, A., *et al.* Strong Candidate for the Breast and Ovarian Cancer Susceptibility Gene BRCA1. *Science* (80-.). **266**, 66–71 (1994).
9. Chen, S. & Parmigiani, G. Meta-Analysis of BRCA1 and BRCA2 Penetrance. *J. Clin. Oncol.* **25**, 1329–1333 (2007).
10. Daly, M. B., Axilbund, J. E., Buys, S., Crawford, B., Farrell, C. D., Friedman, S., Garber, J. E., Goorha, S., Gruber, S. B., Hampel, H., Kaklamani, V., Kohlmann, W., Kurian, A., Litton, J., Marcom, P. K., Nussbaum, R., Offit, K., Pal, T., Pasche, B., *et al.* Genetic/Familial high-risk assessment: Breast and ovarian. *JNCCN J. Natl. Compr. Cancer Netw.* **8**, 562–594 (2010).
11. Couch, F. J., Nathanson, K. L. & Offit, K. Two Decades After BRCA : Setting Care and Prevention. *Science* (80-.). **343**, 1466–1471 (2014).

12. Okano, M., Nomizu, T., Tachibana, K., Nagatsuka, M., Matsuzaki, M., Katagata, N., Ohtake, T., Yokoyama, S., Arai, M. & Nakamura, S. The relationship between BRCA-associated breast cancer and age factors: an analysis of the Japanese HBOC consortium database. *J. Hum. Genet.* **66**, 307–314 (2021).
13. Matulonis, U. A., Sood, A. K., Fallowfield, L., Howitt, B. E., Sehouli, J. & Karlan, B. Y. Ovarian cancer. *Nat. Rev. Dis. Prim.* **2**, 889–908 (2016).
14. Kotsopoulos, J., Gronwald, J., Karlan, B., Rosen, B., Huzarski, T., Moller, P., Lynch, H. T., Singer, C. F., Senter, L., Neuhausen, S. L., Tung, N., Eisen, A., Foulkes, W. D., Ainsworth, P., Sun, P., Lubinski, J. & Narod, S. A. Age-specific ovarian cancer risks among women with a BRCA1 or BRCA2 mutation. *Gynecol. Oncol.* **150**, 85–91 (2018).
15. Lee, A., Moon, B.-I. & Kim, T. H. BRCA1/BRCA2 Pathogenic Variant Breast Cancer: Treatment and Prevention Strategies. *Ann Lab Med* **40**, 114–121 (2020).
16. Tung, N. M. & Garber, J. E. BRCA1/2 testing: Therapeutic implications for breast cancer management. *Br. J. Cancer* **119**, 141–152 (2018).
17. US Food and Drug Administration. FDA approves talazoparib for gBRCAm HER2-negative locally advanced or metastatic breast cancer. (2018). Available at: <https://www.fda.gov/drugs/drug-approvals-and-databases/fda-approves-talazoparib-gbrcam-her2-negative-locally-advanced-or-metastatic-breast-cancer>. (Accessed: 26th May 2021)
18. US Food and Drug Administration. FDA approves rucaparib for maintenance treatment of recurrent ovarian, fallopian tube, or primary peritoneal cancer. (2018). Available at: <https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-rucaparib-maintenance-treatment-recurrent-ovarian-fallopian-tube-or-primary-peritoneal>. (Accessed: 26th May 2021)
19. US Food and Drug Administration. FDA approves olaparib for germline BRCA-mutated metastatic breast cancer. (2018). Available at: <https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-olaparib-germline-brca-mutated-metastatic-breast-cancer>. (Accessed: 26th May 2021)
20. US Food and Drug Administration. FDA approves niraparib for first-line maintenance of advanced ovarian cancer. (2020). Available at: <https://www.fda.gov/drugs/drug-approvals-and-databases/fda-approves-niraparib-first-line-maintenance-advanced-ovarian-cancer>. (Accessed: 26th May 2021)
21. US Food and Drug Administration. FDA approved olaparib (LYNPARZA, AstraZeneca Pharmaceuticals LP) for the maintenance treatment of adult patients with deleterious or suspected deleterious germline or somatic BRCA-mutated (gBRCAm or sBRCAm) advanced epithelial ovarian, fallopian tube or p. (2018). Available at: <https://www.fda.gov/drugs/fda-approved-olaparib-lynparza-astrazeneca-pharmaceuticals-lp-maintenance-treatment-adult-patients>. (Accessed: 26th May 2021)

22. Eastman, A. The formation, isolation and characterization of DNA adducts produced by anticancer platinum complexes. *Pharmacol. Ther.* **34**, 155–166 (1987).
23. Konstantinopoulos, P. A., Ceccaldi, R., Shapiro, G. I. & D'Andrea, A. D. Homologous recombination deficiency: Exploiting the fundamental vulnerability of ovarian cancer. *Cancer Discov.* **5**, 1137–1154 (2015).
24. Galluzzi, L., Senovilla, L., Vitale, I., Michels, J., Martins, I., Kepp, O., Castedo, M. & Kroemer, G. Molecular mechanisms of cisplatin resistance. *Oncogene* **31**, 1869–1883 (2012).
25. Murai, J., Huang, S. N., Das, B. B., Renaud, A., Zhang, Y., Doroshow, J. H., Ji, J., Takeda, S. & Pommier, Y. Trapping of PARP1 and PARP2 by Clinical PARP Inhibitors. *Cancer Res.* **72**, 5588–5599 (2012).
26. Pommier, Y., O'Connor, M. J. & De Bono, J. Laying a trap to kill cancer cells: PARP inhibitors and their mechanisms of action. *Sci. Transl. Med.* **8**, 1–8 (2016).
27. Lord, C. J. & Ashworth, A. PARP inhibitors: Synthetic lethality in the clinic. *Science* (80-.). **355**, 1152–1158 (2017).
28. Konstantinopoulos, P. A., Ceccaldi, R., Shapiro, G. I. & D'andrea, A. D. Homologous Recombination Deficiency: Exploiting the Fundamental Vulnerability of Ovarian Cancer. *Cancer Discov* **5**, 1–18 (2015).
29. Jonsson, P., Bandlamudi, C., Cheng, M. L., Srinivasan, P., Chavan, S. S., Friedman, N. D., Rosen, E. Y., Richards, A. L., Bouvier, N., Selcuklu, S. D., Bielski, C. M., Abida, W., Mandelker, D., Birsoy, O., Zhang, L., Zehir, A., Donoghue, M. T. A., Baselga, J., Offit, K., *et al.* Tumour lineage shapes BRCA-mediated phenotypes. *Nature* **571**, 576–583 (2019).
30. De Leeneer, K., Coene, I., Crombez, B., Simkens, J., Van den Broecke, R., Bols, A., Stragier, B., Vanhoutte, I., De Paepe, A., Poppe, B. & Claes, K. Prevalence of BRCA1/2 mutations in sporadic breast/ovarian cancer patients and identification of a novel de novo BRCA1 mutation in a patient diagnosed with late onset breast and ovarian cancer: implications for genetic testing. *Breast Cancer Res. Treat.* **132**, 87–95 (2012).
31. Maxwell, K. N., Wubbenhorst, B., Wenz, B. M., De Sloover, D., Pluta, J., Emery, L., Barrett, A., Kraya, A. A., Anastopoulos, I. N., Yu, S., Jiang, Y., Chen, H., Zhang, N. R., Hackman, N., D'andrea, K., Daber, R., Morrisette, J. J. D., Mitra, N., Feldman, M., *et al.* BRCA locus-specific loss of heterozygosity in germline BRCA1 and BRCA2 carriers. *Nat. Commun.* **8**, 1–11 (2017).
32. Clarke, C. L., Sandle, J., Jones, A. A., Sofronis, A., Patani, N. R. & Lakhani, S. R. Mapping loss of heterozygosity in normal human breast cells from BRCA1/2 carriers. *Br. J. Cancer* **95**, 515–519 (2006).
33. Cavalli, L. R., Singh, B., Isaacs, C., Dickson, R. B. & Haddad, B. R. Loss of heterozygosity in normal breast epithelial tissue and benign breast lesions in

- BRCA1/2 carriers with breast cancer. *Cancer Genet. Cytogenet.* **149**, 38–43 (2004).
34. Knudson, A. G. Mutation and Cancer: Statistical Study of Retinoblastoma. *Proc. Natl. Acad. Sci.* **68**, 820–823 (1971).
 35. Gudmundsson, J., Johannesdottir, G., Bergthorsson, J. T., Arason, A., Ingvarsson, S., Egilsson, V. & Barkardottir, R. B. Different Tumor Types from BRCA2 Carriers Show Wild-Type Chromosome Deletions on 13q12-q13. *Cancer Res.* **55**, 4830–4832 (1995).
 36. Smith, S. A., Easton, D. F., Evans, D. G. R. & Ponder, B. A. J. Allele losses in the region 17q12-21 in familial breast and ovarian cancer involve the wild-type chromosome. *Nat. Genet.* **2**, 128–131 (1992).
 37. Bryant, H. E., Schultz, N., Thomas, H. D., Parker, K. M., Flower, D., Lopez, E., Kyle, S., Meuth, M., Curtin, N. J. & Helleday, T. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* **434**, 913–917 (2005).
 38. Byrski, T., Huzarski, T., Dent, R., Gronwald, J., Zuziak, D., Cybulski, C., Kladny, J., Gorski, B., Lubinski, J. & Narod, S. A. Response to neoadjuvant therapy with cisplatin in BRCA1-positive breast cancer patients. *Breast Cancer Res. Treat.* **115**, 359–363 (2009).
 39. Byrski, T., Huzarski, T., Dent, R., Gronwald, J., Zuziak, D., Cybulski, C., Kladny, J., Gorski, B., Lubinski, J. & Narod, S. A. Response to neoadjuvant therapy with cisplatin in BRCA1-positive breast cancer patients. *Breast Cancer Res. Treat.* **115**, 359–363 (2009).
 40. Byrski, T., Dent, R., Blecharz, P., Foszczynska-Kloda, M., Gronwald, J., Huzarski, T., Cybulski, C., Marczyk, E., Chrzan, R., Eisen, A., Lubinski, J. & Narod, S. A. Results of a phase II open-label, non-randomized trial of cisplatin chemotherapy in patients with BRCA1-positive metastatic breast cancer. *Breast Cancer Res.* **14**, 1–8 (2012).
 41. Farmer, H., McCabe, H., Lord, C. J., Tutt, A. H. J., Johnson, D. A., Richardson, T. B., Santarosa, M., Dillon, K. J., Hickson, I., Knights, C., Martin, N. M. B., Jackson, S. P., Smith, G. C. M. & Ashworth, A. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917–921 (2005).
 42. Gelmon, K. A., Tischkowitz, M., Mackay, H., Swenerton, K., Robidoux, A., Tonkin, K., Hirte, H., Huntsman, D., Clemons, M., Gilks, B., Yerushalmi, R., Macpherson, E., Carmichael, J. & Oza, A. Olaparib in patients with recurrent high-grade serous or poorly differentiated ovarian carcinoma or triple-negative breast cancer: A phase 2, multicentre, open-label, non-randomised study. *Lancet Oncol.* **12**, 852–861 (2011).
 43. Sandhu, S. K., Schelman, W. R., Wilding, G., Moreno, V., Baird, R. D., Miranda, S., Hylands, L., Riisnaes, R., Forster, M., Omlin, A., Kreischer, N., Thway, K., Gevensleben, H., Sun, L., Loughney, J., Chatterjee, M., Toniatti, C., Carpenter, C.

- L., Iannone, R., *et al.* The poly(ADP-ribose) polymerase inhibitor niraparib (MK4827) in BRCA mutation carriers and patients with sporadic cancer: A phase 1 dose-escalation trial. *Lancet Oncol.* **14**, 882–892 (2013).
44. Balmaña, J., Tung, N. M., Isakoff, S. J., Graña, B., Ryan, P. D., Saura, C., Lowe, E. S., Frewer, P., Winer, E., Baselga, J. & Garber, J. E. Phase I trial of olaparib in combination with cisplatin for the treatment of patients with advanced breast, ovarian and other solid tumors. *Ann. Oncol.* **25**, 1656–1663 (2014).
 45. Gorodnova, T. V., Sokolenko, A. P., Ivantsov, A. O., Iyevleva, A. G., Suspitsin, E. N., Aleksakhina, S. N., Yanus, G. A., Togo, A. V., Maximov, S. Y. & Imyanitov, E. N. High response rates to neoadjuvant platinum-based therapy in ovarian cancer patients carrying germ-line BRCA mutation. *Cancer Lett.* **369**, 363–367 (2015).
 46. Sokolenko, A. P., Savonevich, E. L., Ivantsov, A. O., Raskin, G. A., Kuligina, E. S., Gorodnova, T. V., Preobrazhenskaya, E. V., Kleshchov, M. A., Tiurin, V. I., Mukhina, M. S., Kotiv, K. B., Shulga, A. V., Kuznetsov, S. G., Berlev, I. V. & Imyanitov, E. N. Rapid selection of BRCA1-proficient tumor cells during neoadjuvant therapy for ovarian cancer in BRCA1 mutation carriers. *Cancer Lett.* **397**, 127–132 (2017).
 47. Sokolenko, A. P., Bizin, I. V., Preobrazhenskaya, E. V., Gorodnova, T. V., Ivantsov, A. O., Iyevleva, A. G., Savonevich, E. L., Kotiv, K. B., Kuligina, E. S. & Imyanitov, E. N. Molecular profiles of BRCA1-associated ovarian cancer treated by platinum-based therapy: Analysis of primary, residual and relapsed tumors. *Int. J. Cancer* **146**, 1879–1888 (2020).
 48. Reddy, S. M., Barcenas, C. H., Sinha, A. K., Hsu, L., Moulder, S. L., Tripathy, D., Hortobagyi, G. N. & Valero, V. Long-term survival outcomes of triple-receptor negative breast cancer survivors who are disease free at 5 years and relationship with low hormone receptor positivity. *Br. J. Cancer* **118**, 17–23 (2017).
 49. Goodall, J., Mateo, J., Yuan, W., Mossop, H., Porta, N., Miranda, S., Perez-Lopez, R., Dolling, D., Robinson, D. R., Sandhu, S., Fowler, G., Ebbs, B., Flohr, P., Seed, G., Rodrigues, D. N., Boysen, G., Bertan, C., Atkin, M., Clarke, M., *et al.* Circulating cell-free DNA to guide prostate cancer treatment with PARP inhibition. *Cancer Discov.* **7**, 1006–1017 (2017).
 50. Quigley, D., Alumkal, J. J., Wyatt, A. W., Kothari, V., Foye, A., Lloyd, P., Aggarwal, R., Kim, W., Lu, E., Schwartzman, J., Beja, K., Annala, M., Das, R., Diolaiti, M., Pritchard, C., Thomas, G., Tomlins, S., Knudsen, K., Lord, C. J., *et al.* Analysis of circulating cell-free DnA identifies multiclonal heterogeneity of BRCA2 reversion mutations associated with resistance to PARP inhibitors. *Cancer Discov.* **7**, 999–1005 (2017).
 51. Patch, A. M., Christie, E. L., Etemadmoghadam, D., Garsed, D. W., George, J., Fereday, S., Nones, K., Cowin, P., Alsop, K., Bailey, P. J., Kassahn, K. S., Newell, F., Quinn, M. C. J., Kazakoff, S., Quek, K., Wilhelm-Benartzi, C., Curry, E., Leong, H. S., Hamilton, A., *et al.* Whole-genome characterization of chemoresistant ovarian cancer. *Nature* **521**, 489–494 (2015).

52. Cheng, H. H., Salipante, S. J., Nelson, P. S., Montgomery, B. & Pritchard, C. C. Polyclonal BRCA2 Reversion Mutations Detected in Circulating Tumor DNA After Platinum Chemotherapy in a Patient With Metastatic Prostate Cancer. *JCO Precis. Oncol.* (2018). doi:10.1200/PO
53. Swisher, E. M., Sakai, W., Karlan, B. Y., Wurz, K., Urban, N. & Taniguchi, T. Secondary BRCA1 mutations in BRCA1-mutated ovarian carcinomas with platinum resistance. *Cancer Res.* **68**, 2581–2586 (2008).
54. Norquist, B., Wurz, K. A., Pennil, C. C., Garcia, R., Gross, J., Sakai, W., Karlan, B. Y., Taniguchi, T. & Swisher, E. M. Secondary somatic mutations restoring BRCA1/2 predict chemotherapy resistance in hereditary ovarian carcinomas. *J. Clin. Oncol.* **29**, 3008–3015 (2011).
55. Banda, K., Swisher, E. M., Wu, D., Pritchard, C. C. & Gadi, V. K. Somatic Reversion of Germline *BRCA2* Mutation Confers Resistance to Poly(ADP-ribose) Polymerase Inhibitor Therapy. *JCO Precis. Oncol.* 1–6 (2018). doi:10.1200/PO.17.00044
56. Barber, L. J., Sandhu, S., Chen, L., Campbell, J., Kozarewa, I., Fenwick, K., Assiotis, I., Rodrigues, D. N., Filho, J. S. R., Moreno, V., Mateo, J., Molife, L. R., De Bono, J., Kaye, S., Lord, C. J. & Ashworth, A. Secondary mutations in BRCA2 associated with clinical resistance to a PARP inhibitor. *J. Pathol.* **229**, 422–429 (2013).
57. Lord, C. J. & Ashworth, A. Mechanisms of resistance to therapies targeting BRCA-mutant cancers. *Nature Medicine* 1381–1388 (2013). doi:10.1038/nm.3369
58. Bouwman, P. & Jonkers, J. Molecular pathways: How can BRCA-mutated tumors become resistant to PARP inhibitors? *Clin. Cancer Res.* **20**, 540–547 (2014).
59. Rebbeck, T. R., Mitra, N., Wan, F., Sinilnikova, O. M., Healey, S., McGuffog, L., Chenevix-Trench, G., Easton, D. F., Antoniou, A. C., Nathanson, K. L., Laitman, Y., Kushnir, A., Paluch-Shimon, S., Berger, R., Zidan, J., Friedman, E., Ehrencrona, H., Stenmark-Askmal, M., Einbeigi, Z., *et al.* Association of type and location of BRCA1 and BRCA2 mutations with risk of breast and ovarian cancer. *JAMA - J. Am. Med. Assoc.* **313**, 1347–1361 (2015).
60. Pettitt, S. J., Frankum, J. R., Punta, M., Lise, S., Alexander, J., Chen, Y., Yap, T. A., Haider, S., Tutt, A. N. J. & Lord, C. J. Clinical BRCA1/2 reversion analysis identifies hotspot mutations and predicted neoantigens associated with therapy resistance. *Cancer Discov.* CD-19-1485 (2020). doi:10.1158/2159-8290.cd-19-1485
61. Tutt, A., Tovey, H., Cheang, M. C. U., Kernaghan, S., Kilburn, L., Gazinska, P., Owen, J., Abraham, J., Barrett, S., Barrett-Lee, P., Brown, R., Chan, S., Dowsett, M., Flanagan, J. M., Fox, L., Grigoriadis, A., Gutin, A., Harper-Wynne, C., Hatton, M. Q., *et al.* Carboplatin in BRCA1/2-mutated and triple-negative breast cancer BRCAness subgroups: The TNT Trial. *Nat. Med.* **24**, 628–637 (2018).
62. Wang, Y., Bernhardt, A. J., Cruz, C., Krais, J. J., Nacson, J., Nicolas, E., Peri, S.,

- Gulden, H. Van Der, Heijden, I. Van Der, Brien, S. W. O., Zhang, Y., Harrell, M. I., Johnson, S. F., Candido, F. J., Reis, D., Pharoah, P. D. P., Karlan, B., Gourley, C., Wiest, D. L., *et al.* The BRCA1- D 11q Alternative Splice Isoform Bypasses Germline Mutations and Promotes Therapeutic Resistance to PARP Inhibition and Cisplatin. **76**, 2778–2791 (2016).
63. Wang, Y., Krais, J. J., Bernhardt, A. J., Nicolas, E., Cai, K. Q., Harrell, M. I., Kim, H. H., George, E., Swisher, E. M., Simpkins, F. & Johnson, N. RING domain – deficient BRCA1 promotes PARP inhibitor and platinum resistance. *J. Clin. Invest.* **126**, 3145–3157 (2016).
 64. Drost, R., Dhillon, K. K., Gulden, H. van der, Heijden, I. van der, Brandsma, I., Cruz, C., Chondronasiou, D., Castroviejo-Bermejo, M., Boon, U., Schut, E., Burg, E. van der, Wientjens, E., Pieterse, M., Klijn, C., Klarenbeek, S., Loayza-Puch, F., Elkon, R., Deemter, L. van, Rottenberg, S., *et al.* BRCA1 185delAG tumors may acquire therapy resistance through expression of RING-less BRCA1. *J Clin Invest* **126**, 2903–2918 (2016).
 65. Johnson, N., Johnson, S. F., Yao, W., Li, Y.-C., Choi, Y.-E., Bernhardt, A. J., Wang, Y., Capelletti, M., Sarosiek, K. A., Moreau, L. A., Chowdhury, D., Wickramanayake, A., Harrell, M. I., Liu, J. F., D'andrea, A. D., Miron, A., Swisher, E. M. & Shapiro, G. I. Stabilization of mutant BRCA1 protein confers PARP inhibitor and platinum resistance. *Proc. Natl. Acad. Sci.* **110**, 17041–17046 (2013).
 66. Xu, G., Ross Chapman, J., Brandsma, I., Yuan, J., Mistrik, M., Bouwman, P., Bartkova, J., Gogola, E., Warmerdam, D., Barazas, M., Jaspers, J. E., Watanabe, K., Pieterse, M., Kersbergen, A., Sol, W., Celie, P. H. N., Schouten, P. C., Van Den Broek, B., Salman, A., *et al.* REV7 counteracts DNA double-strand break resection and affects PARP inhibition. *Nature* **521**, 541–544 (2015).
 67. Bunting, S. F., Callén, E., Wong, N., Chen, H.-T., Polato, F., Gunn, A., Bothmer, A., Feldhahn, N., Fernandez-Capetillo, O., Cao, L., Xu, X., Deng, C.-X., Finkel, T., Nussenzweig, M., Stark, J. M. & Nussenzweig, A. 53BP1 Inhibits Homologous Recombination in Brca1-Deficient Cells by Blocking Resection of DNA Breaks. *Cell* **141**, 243–254 (2010).
 68. Bouwman, P., Aly, A., Escandell, J. M., Pieterse, M., Bartkova, J., Van Der Gulden, H., Hiddingh, S., Thanasoula, M., Kulkarni, A., Yang, Q., Haffty, B. G., Tommiska, J., Blomqvist, C., Drapkin, R., Adams, D. J., Nevanlinna, H., Bartek, J., Tarsounas, M., Ganesan, S., *et al.* 53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and BRCA-mutated breast cancers. *Nat. Struct. Mol. Biol.* **17**, 688–695 (2010).
 69. Jaspers, J. E., Kersbergen, A., Boon, U., Sol, W., Van Deemter, L., Zander, S. A., Drost, R., Wientjens, E., Ji, J., Aly, A., Doroshov, J. H., Cranston, A., Martin, N. M. B., Lau, A., O'Connor, M. J., Ganesan, S., Borst, P., Jonkers, J. & Rottenberg, S. Loss of 53BP1 causes PARP inhibitor resistance in BRCA1-mutated mouse mammary tumors. *Cancer Discov.* **3**, 68–81 (2013).

70. Guillemette, S., Serra, R. W., Peng, M., Hayes, J. A., Konstantinopoulos, P. A., Green, M. R., Green, M. R. & Cantor, S. B. Resistance to therapy in BRCA2 mutant cells due to loss of the nucleosome remodeling factor CHD4. *Genes Dev.* **29**, 489–494 (2015).
71. Brady, S. W., McQuerry, J. A., Qiao, Y., Piccolo, S. R., Shrestha, G., Jenkins, D. F., Layer, R. M., Pedersen, B. S., Miller, R. H., Esch, A., Selitsky, S. R., Parker, J. S., Anderson, L. A., Dalley, B. K., Factor, R. E., Reddy, C. B., Boltax, J. P., Li, D. Y., Moos, P. J., *et al.* Combating subclonal evolution of resistant cancer phenotypes. *Nat. Commun.* **8**, (2017).
72. Rottenberg, S., Jaspers, J. E., Kersbergen, A., Van Der Burg, E., Nygren, A. O. H., Zander, S. A. L., Derksen, P. W. B., De Bruin, M., Zevenhoven, J., Lau, A., Boulter, R., Cranston, A., O'Connor, M. J., Martin, N. M. B., Borst, P. & Jonkers, J. High sensitivity of BRCA1-deficient mammary tumors to the PARP inhibitor AZD2281 alone and in combination with platinum drugs. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 17079–17084 (2008).
73. Pettit, S. J., Rehman, F. L., Bajrami, I., Brough, R., Wallberg, F., Kozarewa, I., Fenwick, K., Assiotis, I., Chen, L., Campbell, J., Lord, C. J. & Ashworth, A. A Genetic Screen Using the PiggyBac Transposon in Haploid Cells Identifies Parp1 as a Mediator of Olaparib Toxicity. *PLoS One* **8**, e61520 (2013).
74. Liu, H., Murphy, C. J., Karreth, F. A., Emdal, K. B., Yang, K., White, F. M., Elemento, O., Toker, A., Wulf, G. M. & Cantley, L. C. Identifying and Targeting Sporadic Oncogenic Genetic Aberrations in Mouse Models of Triple-Negative Breast Cancer. *Cancer Discov.* **8**, 354–369 (2017).
75. Shih, H. A., Nathanson, K. L., Seal, S., Collins, N., Stratton, M. R., Rebbeck, T. R. & Weber, B. L. BRCA1 and BRCA2 mutations in breast cancer families with multiple primary cancers. *Clin. Cancer Res.* **6**, 4259–4264 (2000).
76. Labidi-Galy, S. I., Papp, E., Hallberg, D., Niknafs, N., Adleff, V., Noe, M., Bhattacharya, R., Novak, M., Jones, S., Phallen, J., Hruban, C. A., Hirsch, M. S., Lin, D. I., Schwartz, L., Maire, C. L., Tille, J. C., Bowden, M., Ayhan, A., Wood, L. D., *et al.* High grade serous ovarian carcinomas originate in the fallopian tube. *Nat. Commun.* **8**, 1–10 (2017).
77. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
78. Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S. & DePristo, M. A. From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinforma.* **43**, 11.10.1–11.10.33 (2013).
79. Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras, T. R. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

80. Kurnit, K. C., Fleming, G. F. & Lengyel, E. Updates and New Options in Advanced Epithelial Ovarian Cancer Treatment. *Obstet. Gynecol.* **137**, 108–129 (2021).
81. Venkitaraman, A. R. Cancer Susceptibility and the Functions of BRCA1 and BRCA2. *Cell* **108**, 171–182 (2002).
82. Moynahan, M. E., Pierce, A. J. & Jasin, M. BRCA2 Is Required for Homology-Directed Repair of Chromosomal Breaks. *Mol. Cell* **7**, 263–272 (2001).
83. Lawrence, M. S., Stojanov, P., Polak, P., Kryukov, G. V., Cibulskis, K., Sivachenko, A., Carter, S. L., Stewart, C., Mermel, C. H., Roberts, S. A., Kiezun, A., Hammerman, P. S., McKenna, A., Drier, Y., Zou, L., Ramos, A. H., Pugh, T. J., Stransky, N., Helman, E., *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
84. Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A. J. R., Behjati, S., Biankin, A. V., Bignell, G. R., Bolli, N., Borg, A., Børresen-Dale, A. L., Boyault, S., Burkhardt, B., Butler, A. P., Caldas, C., Davies, H. R., Desmedt, C., Eils, R., Eyfjörd, J. E., Foekens, J. A., *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
85. Wu, H.-X., Wang, Z.-X., Zhao, Q., Chen, D.-L., He, M.-M., Yang, L.-P., Wang, Y.-N., Jin, Y., Ren, C., Luo, H.-Y., Wang, Z.-Q. & Wang, F. Tumor mutational and indel burden: a systematic pan-cancer evaluation as prognostic biomarkers. *Ann. Transl. Med.* **7**, 640–640 (2019).
86. Lai, Z., Markovets, A., Ahdesmaki, M., Chapman, B., Hofmann, O., Mcewen, R., Johnson, J., Dougherty, B., Barrett, J. C. & Dry, J. R. VarDict: A novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res.* **44**, 1–11 (2016).
87. Kim, S., Scheffler, K., Halpern, A. L., Bekritsky, M. A., Noh, E., Källberg, M., Chen, X., Kim, Y., Beyter, D., Krusche, P. & Saunders, C. T. Strelka2: fast and accurate calling of germline and somatic variants. *Nat. Methods* **15**, 591–594 (2018).
88. Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E. S. & Getz, G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
89. Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., Miller, C. A., Mardis, E. R., Ding, L. & Wilson, R. K. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
90. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, 1–7 (2010).
91. Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B. S. & Swanton, C.

- deconstructSigs: Delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol.* **17**, 31 (2016).
92. Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., Karyadi, D., Cannon-Albright, L. A., Teerlink, C. C., Stanford, J. L., Isaacs, W. B., Xu, J., Cooney, K. A., Lange, E. M., Schleutker, J., Carpten, J. D., *et al.* REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am. J. Hum. Genet.* **99**, 877–885 (2016).
 93. Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S. & Mesirov, J. P. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–50 (2005).
 94. Schwarz, J. M., Cooper, D. N., Schuelke, M. & Seelow, D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods* **11**, 361–362 (2014).
 95. Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. & Mesirov, J. P. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
 96. Carpten, J. D., Faber, A. L., Horn, C., Donoho, G. P., Briggs, S. L., Robbins, C. M., Hostetter, G., Boguslawski, S., Moses, T. Y., Savage, S., Uhlik, M., Lin, A., Du, J., Qian, Y.-W., Zeckner, D. J., Tucker-Kellogg, G., Touchman, J., Patel, K., Mousses, S., *et al.* A transforming mutation in the pleckstrin homology domain of AKT1 in cancer. *Nature* **448**, 439–445 (2007).
 97. Chen, Y., Huang, L., Dong, Y., Tao, C., Zhang, R., Shao, H. & Shen, H. Effect of AKT1 (p. E17K) Hotspot Mutation on Malignant Tumorigenesis and Prognosis. *Frontiers in Cell and Developmental Biology* **8**, 1–12 (2020).
 98. Ruault, M., van der Bruggen, P., Brun, M. E., Boyle, S., Roizès, G. & De Sario, A. New BAGE (B melanoma antigen) genes mapping to the juxtacentromeric regions of human chromosomes 13 and 21 have a cancer/testis expression profile. *Eur. J. Hum. Genet.* **10**, 833–840 (2002).
 99. Jaspers, J. E., Kersbergen, A., Boon, U., Sol, W., Van Deemter, L., Zander, S. A., Drost, R., Wientjens, E., Ji, J., Aly, A., Doroshov, J. H., Cranston, A., Martin, N. M. B., Lau, A., O'Connor, M. J., Ganesan, S., Borst, P., Jonkers, J. & Rottenberg, S. Loss of 53BP1 causes PARP inhibitor resistance in BRCA1-mutated mouse mammary tumors. *Cancer Discov.* **3**, 68–81 (2013).
 100. Inagaki-Kawata, Y., Yoshida, K., Kawaguchi-Sakita, N., Kawashima, M., Nishimura, T., Senda, N., Shiozawa, Y., Takeuchi, Y., Inoue, Y., Sato-Otsubo, A., Fujii, Y., Nannya, Y., Suzuki, E., Takada, M., Tanaka, H., Shiraishi, Y., Chiba, K., Kataoka, Y., Torii, M., *et al.* Genetic and clinical landscape of breast cancers with germline BRCA1/2 variants. *Commun. Biol.* **3**, 1–9 (2020).

101. Boyarskikh, U. A., Gulyaeva, L. F., Avdalyan, A. M., Kechin, A. A., Khrapov, E. A., Lazareva, D. G., Kushlinskii, N. E., Melkonyan, A., Arakelyan, A. & Filipenko, M. L. Spectrum of TP53 Mutations in BRCA1/2 Associated High-Grade Serous Ovarian Cancer. *Front. Oncol.* **10**, 1103 (2020).
102. Holstege, H., Joosse, S. A., Th van Oostrom, C. M., Nederlof, P. M., de Vries, A. & Jonkers, J. High Incidence of Protein-Truncating TP53 Mutations in BRCA1-Related Breast Cancer. *Cancer Res* **69**, 3625–3658 (2009).
103. Meric-Bernstam, F., Frampton, G. M., Ferrer-Lozano, J., Yelensky, R., Perez-Fidalgo, J. A., Wang, Y., Palmer, G. A., Ross, J. S., Miller, V. A., Su, X., Eroles, P., Barrera, J. A., Burgues, O., Lluch, A. M., Zheng, X., Sahin, A., Stephens, P. J., Mills, G. B., Cronin, M. T., *et al.* Concordance of Genomic Alterations between Primary and Recurrent Breast Cancer. *Mol. Cancer Ther.* **13**, 1382–1389 (2014).
104. Masoodi, T., Siraj, S., Siraj, A. K., Azam, S., Qadri, Z., Parvathareddy, S. K., Tulbah, A., Al-Dayel, F., Alhusaini, H., Alomar, O., Al-Badawi, I. A., Alkuraya, F. S. & Al-Kuraya, K. S. Genetic heterogeneity and evolutionary history of high-grade ovarian carcinoma and matched distant metastases. *Br. J. Cancer* **122**, 1219–1230 (2020).
105. Kim, Y.-M., Lee, S.-W., Chun, S.-M., Kim, D.-Y., Kim, J.-H., Kim, K.-R., Kim, Y.-T., Nam, J.-H., Van Hummelen, P., Macconail, L. E., Hahn, W. C. & Jang, J. Analysis and Comparison of Somatic Mutations in Paired Primary and Recurrent Epithelial Ovarian Cancer Samples. *PLoS One* **9**, e99451 (2014).
106. Brose, M. S., Rebbeck, T. R., Calzone, K. A., Stopfer, J. E., Nathanson, K. L. & Weber, B. L. Cancer risk estimates for BRCA1 mutation carriers identified in a risk evaluation program. *J. Natl. Cancer Inst.* **94**, 1365–1372 (2002).
107. Telli, M. L., Timms, K., Reid, J., Hennessy, B., Mills, G., Jensen, K. C., Barry, W. T., Winer, E. P., Tung, N. M., Isakoff, S. J., Ryan, P., Melinda, L. T., Kirsten, M. T., Julia, R., Bryan, H., Gordon, B. M., Kristin, C. J., Zoltan, S., William, T. B., *et al.* Homologous Recombination Deficiency (HRD) Score Predicts Response to Platinum-Containing Neoadjuvant Chemotherapy in Patients with Triple Negative Breast Cancer. *Clin. Cancer Res.* **22**, 3764–3773 (2016).
108. Marquard, A. M., Eklund, A. C., Joshi, T., Krzystanek, M., Favero, F., Wang, Z. C., Richardson, A. L., Silver, D. P., Szallasi, Z. & Birkbak, N. J. Pan-cancer analysis of genomic scar signatures associated with homologous recombination deficiency suggests novel indications for existing cancer drugs. *Biomark. Res.* **3**, 9 (2015).
109. B A da Costa, A. A., do Canto, L. M., Jonas Larsen, S., Regina Gonçalves Ribeiro, A., Eduardo Stecca, C., Høgh Petersen, A., Aagaard, M. M., de Brot, L., Baumbach, J., Baiocchi, G., Isabel Achatz, M. & Regina Rogatto, S. Genomic profiling in ovarian cancer retreated with platinum based chemotherapy presented homologous recombination deficiency and copy number imbalances of CCNE1 and RB1 genes. *BMC Cancer* **19**, (2019).
110. Abkevich, V., Timms, K. M., Hennessy, B. T., Potter, J., Carey, M. S., Meyer, L. A., Smith-McCune, K., Broaddus, R., Lu, K. H., Chen, J., Tran, T. V., Williams, D.,

- Iliev, D., Jammulapati, S., Fitzgerald, L. M., Krivak, T., Deloia, J. A., Gutin, A., Mills, G. B., *et al.* Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. *Br. J. Cancer* **107**, 1776–1782 (2012).
111. Birkbak, N. J., Wang, Z. C., Kim, J.-Y., Eklund, A. C., Li, Q., Tian, R., Bowman-Colin, C., Li, Y., Greene-Colozzi, A., Iglehart, J. D., Tung, N., Ryan, P. D., Garber, J. E., Silver, D. P., Szallasi, Z. & Richardson, A. L. Telomeric Allelic Imbalance Indicates Defective DNA Repair and Sensitivity to DNA-Damaging Agents. *Cancer Discov.* 366–375 (2012). doi:10.1158/2159
 112. Popova, T., Manié, E., Rieunier, G., Caux-Moncoutier, V., Tirapo, C., Dubois, T., Delattre, O., Sigal-Zafrani, B., Bollet, M., Longy, M., Houdayer, C., Sastre-Garau, X., Vincent-Salomon, A., Stoppa-Lyonnet, D. & Stern, M. H. Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. *Cancer Res.* **72**, 5454–5462 (2012).
 113. Favero, F., Joshi, T., Marquard, A. M., Birkbak, N. J., Krzystanek, M., Li, Q., Szallasi, Z. & Eklund, A. K. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann. Oncol.* **26**, 64–70 (2014).
 114. Mermel, C. H., Schumacher, S. E., Hill, B., Meyerson, M. L., Beroukhi, R. & Getz, G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **12**, R41 (2011).
 115. Kraya, A. A., Maxwell, K. N., Wubbenhorst, B., Wenz, B. M., Pluta, J., Rech, A. J., Dorfman, L. M., Lunceford, N., Barrett, A., Mitra, N., Morrisette, J. J. D., Feldman, M., Nayak, A., Domchek, S. M., Vonderheide, R. H. & Nathanson, K. L. Genomic Signatures Predict the Immunogenicity of BRCA-Deficient Breast Cancer. (2019). doi:10.1158/1078-0432.CCR-18-0468
 116. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
 117. Ronique Geoffroy, V., Herenger, Y., Kress, A., Stoetzel, C., Lie Piton, A., Lè Ne Dollfus, H. & Muller, J. AnnotSV: an integrated tool for structural variations annotation. *Bioinformatics* **34**, 3572–3574 (2018).
 118. Mehta, G. A., Parker, J. S., Silva, G. O., Hoadley, K. A., Perou, C. M. & Gatz, M. L. Amplification of SOX4 promotes PI3K/Akt signaling in human breast cancer. *Breast Cancer Res. Treat.* **162**, 439–450 (2017).
 119. Marmorstein, R. & Zhou, M. Writers and Readers of Histone Acetylation : Structure, Mechanism, and Inhibition. *Cold Spring Harb Perspect Biol* **6**, 1–26 (2014).
 120. Li, Y. & Seto, E. HDACs and HDAC inhibitors in cancer development and therapy. *Cold Spring Harb. Perspect. Med.* **6**, 1–34 (2016).
 121. Chen, L. F. Tumor suppressor function of RUNX3 in breast cancer. *J. Cell.*

Biochem. **113**, 1470–1477 (2012).

122. Yasui, K., Mihara, S., Zhao, C., Okamoto, H., Saito-Ohara, F., Tomida, A., Funato, T., Yokomizo, A., Naito, S., Imoto, I., Tsuruo, T. & Inazawa, J. Alteration in Copy Numbers of Genes as a Mechanism for Acquired Drug Resistance. *Cancer Res.* **64**, 1403–1410 (2004).
123. Patel, J. N., Braicu, I., Timms, K. M., Solimeno, C., Tshiaba, P., Reid, J., Lanchbury, J. S., Darb-Esfahani, S., Ganapathi, M. K., Sehouli, J. & Ganapathi, R. N. Characterisation of homologous recombination deficiency in paired primary and recurrent high grade serous ovarian cancer. *Br. J. Cancer* 1–7 (2018). doi:10.1038/s41416-018-0268-6
124. Grushko, T. A., Dignam, J. J., Das, S., Blackwood, A. M., Perou, C. M., Ridderstråle, K. K., Anderson, K. N., Wei, M. J., Adams, A. J., Hagos, F. G., Sveen, L., Lynch, H. T., Weber, B. L. & Olopade, O. I. MYC Is Amplified in BRCA1-Associated Breast Cancers. *Clin. Cancer Res.* **10**, 499–507 (2004).
125. Castro, E., Jugurnauth-Little, S., Karlsson, Q., Al-Shahrour, F., Piñeiro-Yañez, E., Van de Poll, F., Leongamornlert, D., Dadaev, T., Govindasami, K., Guy, M., Eeles, R. & Kote-Jarai, Z. High burden of copy number alterations and c-MYC amplification in prostate cancer from BRCA2 germline mutation carriers. *Ann. Oncol.* **26**, 2293–2300 (2015).
126. Taylor, R. A., Fraser, M., Livingstone, J., Melijah, S., Espiritu, G., Thorne, H., Huang, V., Lo, W., Shiah, Y.-J., Yamaguchi, T. N., Sliwinski, A., Horsburgh, S., Meng, A., Heisler, L. E., Yu, N., Yousif, F., Papargiris, M., Lawrence, M. G., Timms, L., *et al.* Germline BRCA2 mutations drive prostate cancers with distinct evolutionary trajectories. *Nat. Commun.* **8**, (2017).
127. Hoadley, K. A., Yau, C., Hinoue, T., Wolf, D. M., Lazar, A. J., Drill, E., Shen, R., Taylor, A. M., Cherniack, A. D., Thorsson, V., Akbani, R., Bowlby, R., Wong, C. K., Wiznerowicz, M., Sanchez-Vega, F., Robertson, A. G., Schneider, B. G., Lawrence, M. S., Noushmehr, H., *et al.* Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* **173**, 291-304.e6 (2018).
128. Schuierer, S., Carbone, W., Knehr, J., Petitjean, V., Fernandez, A., Sultan, M. & Roma, G. A comprehensive assessment of RNA-seq protocols for degraded and low-quantity samples. *BMC Genomics* **18**, (2017).
129. Cieslik, M., Chugh, R., Wu, Y. M., Wu, M., Brennan, C., Lonigro, R., Su, F., Wang, R., Siddiqui, J., Mehra, R., Cao, X., Lucas, D., Chinnaiyan, A. M. & Robinson, D. The use of exome capture RNA-seq for highly degraded RNA with application to clinical cancer sequencing. *Genome Res.* **25**, 1372–1381 (2015).
130. Mercer, T. R., Clark, M. B., Crawford, J., Brunck, M. E., Gerhardt, D. J., Taft, R. J., Nielsen, L. K., Dinger, M. E. & Mattick, J. S. Targeted sequencing for gene discovery and quantification using RNA CaptureSeq. *Nat. Protoc.* **9**, 989–1009 (2014).

131. Mercer, T. R., Gerhardt, D. J., Dinger, M. E., Crawford, J., Trapnell, C., Jeddelloh, J. A., Mattick, J. S. & Rinn, J. L. Targeted RNA sequencing reveals the deep complexity of the human transcriptome. *Nat. Biotechnol.* **30**, 99–104 (2011).
132. Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. & Salzberg, S. L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, 1650–1667 (2016).
133. Sonesson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences. *F1000Research* **4**, 1521 (2016).
134. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
135. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, (2010).
136. Kwon, A. T., Arenillas, D. J., Hunt, R. W. & Wasserman, W. W. oPOSSUM-3: Advanced Analysis of Regulatory Motif Over-Representation Across Genes or ChIP-Seq Datasets. *G3* **2**, 987–1002 (2012).
137. Haas, B. J., Dobin, A., Li, B., Stransky, N., Pochet, N. & Regev, A. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol.* **20**, (2019).
138. Pai, V. C., Hsu, C. C., Chan, T. S., Liao, W. Y., Chuu, C. P., Chen, W. Y., Li, C. R., Lin, C. Y., Huang, S. P., Chen, L. T. & Tsai, K. K. ASPM promotes prostate cancer stemness and progression by augmenting Wnt–Dvl-3– β -catenin signaling. *Oncogene* **38**, 1340–1353 (2019).
139. Brüning-Richardson, A., Bond, J., Alsiary, R., Richardson, J., Cairns, D. A., McCormack, L., Hutson, R., Burns, P., Wilkinson, N., Hall, G. D., Morrison, E. E. & Bell, S. M. ASPM and microcephalin expression in epithelial ovarian cancer correlates with tumour grade and survival. *Br. J. Cancer* **104**, 1602–1610 (2011).
140. Xu, Z., Zhang, Q. I., Luh, F., Jin, B. & Liu, X. Overexpression of the ASPM gene is associated with aggressiveness and poor outcome in bladder cancer. *Oncol. Lett.* **17**, 1865–1876 (2019).
141. Wheatley, S. P. & Altieri, D. C. Survivin at a glance. *J. Cell Sci.* **132**, (2019).
142. Liu, H. Y. & Zhang, C. J. Identification of differentially expressed genes and their upstream regulators in colorectal cancer. *Cancer Gene Ther.* **24**, 244–250 (2017).
143. Ma, Y., Yang, Y., Wang, F., Wei, Q. & Qin, H. Hippo-YAP signaling pathway: A new paradigm for cancer therapy. *Int. J. Cancer* **137**, 2275–2286 (2015).
144. Hulin, J.-A., Tommasi, S., Elliot, D., Gui Hu, D., Lewis, B. C. & Mangoni, A. A. MiR-193b regulates breast cancer cell migration and vasculogenic mimicry by

- targeting dimethylarginine dimethylaminohydrolase 1 OPEN. *Sci. Rep.* **7**, (2017).
145. Zhang, J., Qin, J. & Su, Y. miR-193b-3p possesses anti-tumor activity in ovarian carcinoma cells by targeting p21-activated kinase 3. *Biomed. Pharmacother.* **96**, 1275–1282 (2017).
 146. Choi, Y. E., Pan, Y., Park, E., Konstantinopoulos, P. A., De, S., D'Andrea, A. D. & Chowdhury, D. MicroRNAs down-regulate homologous recombination in the G1 phase of cycling cells to maintain genomic stability. *Elife* **2014**, (2014).
 147. Santilli, G., Cervellera, M. N., Johnson, T. K., Lewis, R. E., Iacobelli, S. & Sala, A. PARP co-activates B-MYB through enhanced phosphorylation at cyclin/cdk2 sites. *Oncogene* **20**, 8167–8174 (2001).
 148. Bieche, I., Pennaneach, V., Driouch, K., Vacher, S., Zaremba, T., Susini, A., Lidereau, R. & Hall, J. Variations in the mRNA expression of poly(ADP-ribose) polymerases, poly(ADP-ribose) glycohydrolase and ADP-ribosylhydrolase 3 in breast tumors and impact on clinical outcome. *Int. J. Cancer* **133**, 2791–2800 (2013).
 149. Eismann, J., Heng, Y. J., Waldschmidt, J. M., Vlachos, I. S., Gray, K. P., Matulonis, U. A., Konstantinopoulos, P. A., Charles, ·, Murphy, J., Nabavi, S., Gerburg, · & Wulf, M. Transcriptome analysis reveals overlap in fusion genes in a phase I clinical cohort of TNBC and HGSOC patients treated with buparlisib and olaparib. *J. Cancer Res. Clin. Oncol.* **146**, 503–514 (2020).
 150. Li, Z.-X., Zhu, Q.-N., Zhang, H.-B., Hu, Y., Wang, G. & Zhu, Y.-S. MALAT1: a potential biomarker in cancer. *Cancer Manag. Res.* 6757–6768 (2018). doi:10.2147/CMAR.S169406
 151. Ossovskaya, V., Koo, I. C., Kaldjian, E. P., Alvares, C. & Sherman, B. M. Upregulation of poly (ADP-Ribose) polymerase-1 (PARP1) in triple-negative breast cancer and other primary human tumor types. *Genes and Cancer* **1**, 812–821 (2010).
 152. Bi, F. F., Li, D. & Yang, Q. Hypomethylation of ETS transcription factor binding sites and upregulation of PARP1 expression in endometrial cancer. *Biomed Res. Int.* **2013**, 1–5 (2013).
 153. Lheureux, S., Bruce, J. P., Burnier, J. V., Karakasis, K., Shaw, P. A., Clarke, B. A., Yang, S. Y. C., Quevedo, R., Li, T., Dowar, M., Bowering, V., Pugh, T. J. & Oza, A. M. Somatic BRCA1/2 recovery as a resistance mechanism after exceptional response to poly (ADP-ribose) polymerase inhibition. *J. Clin. Oncol.* **35**, 1240–1249 (2017).
 154. Rojo, F., García-Parra, J., Zazo, S., Tusquets, I., Ferrer-Lozano, J., Menendez, S., Eroles, P., Chamizo, C., Servitja, S., Ramírez-Merino, N., Lobo, F., Bellosillo, B., Corominas, J. M., Yelamos, J., Serrano, S., Lluch, A., Rovira, A. & Albanell, J. Nuclear PARP-1 protein overexpression is associated with poor overall survival in early breast cancer. *Ann. Oncol.* **23**, 1156–1164 (2012).

155. Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40**, 1413–1415 (2008).
156. Zhang, Y., Qian, J., Gu, C. & Yang, Y. Alternative splicing and cancer: a systematic review. *Signal Transduct. Target. Ther.* **6**, (2021).
157. Oltean, S. & Bates, D. O. Hallmarks of alternative splicing in cancer. *Oncogene* **33**, 5311–5318 (2014).
158. Bonnal, S. C., López-Oreja, I. & Valcárcel, J. Roles and mechanisms of alternative splicing in cancer — implications for care. *Nat. Rev. Clin. Oncol.* **17**, 457–474 (2020).
159. Rhine, C. L., Cygan, K. J., Soemedi, R., Maguire, S., Murray, M. F., Monaghan, S. F. & Fairbrother, W. G. Hereditary cancer genes are highly susceptible to splicing mutations. *PLoS Genet.* **14**, (2018).
160. Vitting-Seerup, K. & Sandelin, A. The landscape of isoform switches in human cancers. *Mol. Cancer Res.* **15**, 1206–1220 (2017).
161. Vitting-Seerup, K., Porse, B. T., Sandelin, A. & Waage, J. SpliceR: An R package for classification of alternative splicing and prediction of coding potential from RNA-seq data. *BMC Bioinformatics* **15**, 81 (2014).
162. Mukherjee, N., Corcoran, D. L., Nusbaum, J. D., Reid, D. W., Georgiev, S., Hafner, M., Ascano, M., Tuschl, T., Ohler, U. & Keene, J. D. Integrative Regulatory Mapping Indicates that the RNA-Binding Protein HuR Couples Pre-mRNA Processing and mRNA Stability. *Mol. Cell* **43**, 327–339 (2011).
163. Fatscher, T., Boehm, V., Weiche, B. & Gehring, N. H. The interaction of cytoplasmic poly(A)-binding protein with eukaryotic initiation factor 4G suppresses nonsense-mediated mRNA decay. *RNA* **20**, 1579–1592 (2014).
164. Behm-Ansmant, I., Gatfield, D., Rehwinkel, J., Hilgers, V. & Izaurralde, E. A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. *EMBO J.* **26**, 1591–1601 (2007).
165. Wang, J., Guo, Y., Chu, H., Guan, Y., Bi, J. & Wang, B. Multiple Functions of the RNA-Binding Protein HuR in Cancer Progression, Treatment Responses and Prognosis. *Int. J. Mol. Sci* **14**, 10015–10041 (2013).
166. Mangus, D. A., Evans, M. C. & Jacobson, A. Poly(A)-binding proteins: Multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biology* **4**, 223 (2003).
167. Rebbeck, T. R., Friebel, T. M., Friedman, E., Hamann, U., Huo, D., Kwong, A., Olah, E., Olopade, O. I., Solano, A. R., Teo, S. H., Thomassen, M., Weitzel, J. N., Chan, T. L., Couch, F. J., Goldgar, D. E., Kruse, T. A., Palmero, E. I., Park, S. K., Torres, D., *et al.* Mutational spectrum in a worldwide study of 29,700 families with BRCA1 or BRCA2 mutations. *Hum. Mutat.* **39**, 593–620 (2018).

168. Sakai, W., Swisher, E. M., Karlan, B. Y., Agarwal, M. K., Higgins, J., Friedman, C., Villegas, E., Jacquemont, C., Farrugia, D. J., Couch, F. J., Urban, N. & Taniguchi, T. Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. *Nature* **451**, 1116–1120 (2008).
169. Sakai, W., Swisher, E. M., Jacquemont, C., Chandramohan, K. V., Couch, F. J., Langdon, S. P., Wurz, K., Higgins, J., Villegas, E. & Taniguchi, T. Functional restoration of BRCA2 protein by secondary BRCA2 mutations in BRCA2-mutated ovarian carcinoma. *Cancer Res.* **69**, 6381–6386 (2009).
170. Park, P. H., Yamamoto, T. M., Li, H., Alcivar, A. L., Xia, B., Wang, Y., Bernhardt, A. J., Turner, K. M., Kossenkova, A. V., Watson, Z. L., Behbakht, K., Casadei, S., Swisher, E. M., Mischel, P. S., Johnson, N. & Bitler, B. G. Amplification of the Mutation-Carrying BRCA2 Allele Promotes RAD51 Loading and PARP Inhibitor Resistance in the Absence of Reversion Mutations. *Mol. Cancer Ther.* **19**, 602–613 (2020).
171. Lin, K. K., Harrell, M. I., Oza, A. M., Oaknin, A., Ray-Coquard, I., Tinker, A. V., Helman, E., Radke, M. R., Say, C., Vo, L.-T., Mann, E., Isaacson, J. D., Maloney, L., O'Malley, D. M., Chambers, S. K., Kaufmann, S. H., Scott, C. L., Konecny, G. E., Coleman, R. L., *et al.* BRCA Reversion Mutations in Circulating Tumor DNA Predict Primary and Acquired Resistance to the PARP Inhibitor Rucaparib in High-Grade Ovarian Carcinoma. *Cancer Discov.* CD-18-0715 (2018). doi:10.1158/2159-8290.CD-18-0715
172. Feng, Z., Scott, S. P., Bussen, W., Sharma, G. G., Guo, G., Pandita, T. K. & Powell, S. N. Rad52 inactivation is synthetically lethal with BRCA2 deficiency. *PNAS* **108**, 686–691 (2010).
173. Bouwman, P. & Jonkers, J. The effects of deregulated DNA damage signalling on cancer chemotherapy response and resistance. *Nature Reviews Cancer* **12**, 587–598 (2012).
174. Broustas, C. G., Lieberman & Howard B. DNA Damage Response Genes and the Development of Cancer Metastasis. *Radiat. Res.* **181**, 111–130 (2014).
175. O'Connor, M. J. Targeting the DNA Damage Response in Cancer. *Molecular Cell* **60**, 547–560 (2015).
176. Maréchal, A. & Zou, L. DNA Damage Sensing by the ATM and ATR Kinases. *Cold Spring Harb. Perspect. Biol.* **5**, (2013).
177. Kim, N.-G., Choi, Y. R., Baek, M. J., Kim, Y. H., Kang, H., Kim, N. K., Min, J. S. & Kim, H. Frameshift Mutations at Coding Mononucleotide Repeats of the hRAD50 Gene in Gastrointestinal Carcinomas with Microsatellite Instability 1. (2001).
178. Menoyo, A., Alazzouzi, H., Espin, E., Armengol, M., Yamamoto, H. & Schwartz, S. Somatic Mutations in the DNA Damage-Response Genes ATR and CHK1 in Sporadic Stomach Tumors with Microsatellite Instability. *Cancer Res.* **61**, 7727–7730 (2001).

179. Chen, H.-Y., Shao, C.-J., Chen, F.-R., Kwan, A.-L. & Chen, Z.-P. Role of ERCC1 promoter hypermethylation in drug resistance to cisplatin in human gliomas. *Int. J. Cancer* **126**, 1944–1954 (2010).
180. Van Allen, E. M., Mouw, K. W., Kim, P., Iyer, G., Wagle, N., Al-Ahmadie, H., Zhu, C., Ostrovnaya, I., Kryukov, G. V., O'connor, K. W., Sfakianos, J., Garcia-Grossman, I., Kim, J., Guancial, E. A., Bambury, R., Bahl, S., Gupta, N., Farlow, D., Qu, A., *et al.* Somatic ERCC2 mutations correlate with cisplatin sensitivity in muscle-invasive urothelial carcinoma. *Cancer Discov.* **4**, 1140–1153 (2014).
181. Yadav, S., Anbalagan, M., Baddoo, M., Chellamuthu, V. K., Mukhopadhyay, S., Woods, C., Jiang, W., Moroz, K., Flemington, E. K. & Makridakis, N. Somatic mutations in the DNA repairome in prostate cancers in African Americans and Caucasians. *Oncogene* **39**, 4299–4311 (2020).
182. Cunningham, J. M., Kim, C. Y., Christensen, E. R., Tester, D. J., Parc, Y., Burgart, L. J., Halling, K. C., McDonnell, S. K., Schaid, D. J., Walsh Vockley, C., Kubly, V., Nelson, H., Michels, V. V. & Thibodeau, S. N. The frequency of hereditary defective mismatch repair in a prospective series of unselected colorectal carcinomas. *Am. J. Hum. Genet.* **69**, 780–790 (2001).
183. Bian, L., Meng, Y., Zhang, M. & Li, D. MRE11-RAD50-NBS1 complex alterations and DNA damage response: Implications for cancer treatment. *Molecular Cancer* **18**, 1–14 (2019).
184. Wang, C., Horiuchi, A., Imai, T., Ohira, S., Itoh, K., Nikaido, T., Katsuyama, Y. & Konishi, I. Expression of BRCA1 protein in benign, borderline, and malignant epithelial ovarian neoplasms and its relationship to methylation and allelic loss of the BRCA1 gene. *J. Pathol.* **202**, 215–223 (2004).
185. Shilpa, V., Bhagat, R., Premalata, C. S., Pallavi, V. R., Ramesh, G. & Krishnamoorthy, L. BRCA1 promoter hypermethylation and protein expression in ovarian carcinoma - An Indian study. *Tumor Biol.* **35**, 4277–4284 (2014).
186. Garg, K., Levine, D. A., Olvera, N., Dao, F., Bisogna, M., Secord, A. A., Berchuck, A., Cerami, E., Schultz, N. & Soslow, R. A. BRCA1 immunohistochemistry in a molecularly characterized cohort of ovarian high-grade serous carcinomas. *Am. J. Surg. Pathol.* **37**, 138–146 (2013).
187. Goltsev, Y., Samusik, N., Kennedy-Darling, J., Bhate, S., Hale, M., Vazquez, G., Black, S. & Nolan, G. P. Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968-981.e15 (2018).
188. LoRusso, P., Pilat, M. J. P., Santa-Maria, C. A., Connolly, R. M., Roesch, E. E., Afghahi, A., Han, H. S., Nanda, R., Wulf, G. M., Assad, H., Park, H., Dees, E. C., Force, J. M., Noonan, A. M., Brufsky, A., Abramson, V. G., Haley, B. B., Buys, S. S., Sharon, E., *et al.* Trial in progress: A phase II open-label, randomized study of PARP inhibition (olaparib) either alone or in combination with anti-PD-L1 therapy (atezolizumab) in homologous DNA repair (HDR) deficient, locally advanced or metastatic non-HER2-positive breast cancer. *J. Clin. Oncol.* **38**, TPS1102–TPS1102 (2020).

189. Zhang, J., Shih, D. J. H. & Lin, S. Y. Role of DNA repair defects in predicting immunotherapy response. *Biomark. Res.* **8**, 1–8 (2020).
190. Zhang, G., Liu, C., Bai, H., Cao, G., Cui, R. & Zhang, Z. Combinatorial therapy of immune checkpoint and cancer pathways provides a novel perspective on ovarian cancer treatment (Review). *Oncol. Lett.* **17**, 2583–2591 (2019).
191. Matulonis, U. A., Shapira-Frommer, R., Santin, A. D., Lisyanskaya, A. S., Pignata, S., Vergote, I., Raspagliesi, F., Sonke, G. S., Birrer, M., Provencher, D. M., Sehouli, J., Colombo, N., González-Martín, A., Oaknin, A., Ottevanger, P. B., Rudaitis, V., Katchar, K., Wu, H., Keefe, S., *et al.* Antitumor activity and safety of pembrolizumab in patients with advanced recurrent ovarian cancer: results from the phase II KEYNOTE-100 study. *Ann. Oncol.* **30**, 1080–1087 (2019).
192. Disis, M. L., Taylor, M. H., Kelly, K., Thaddeus Beck, J., Gordon, M., Moore, K. M., Patel, M. R., Chaves, J., Park, H., Mita, A. C., Hamilton, E. P., Annunziata, C. M., Juergen Grote, H., von Heydebreck, A., Grewal, J., Chand, V. & Gulley, J. L. Efficacy and Safety of Avelumab for Patients With Recurrent or Refractory Ovarian Cancer Phase 1b: Results From the JAVELIN Solid Tumor Trial. *JAMA Oncol.* **5**, 393–401 (2019).
193. Domchek, S. M., Postel-Vinay, S., Im, S.-A., Park, Y. H., Delord, J.-P., Italiano, A., Alexandre, J., You, B., Bastian, S., Krebs, M. G., Wang, D., Waqar, S. N., Lanasa, M., Rhee, J., Gao, H., Rocher-Ros, V., Jones, E. V., Gulati, S., Coenen-Stass, A., *et al.* Olaparib and durvalumab in patients with germline BRCA-mutated metastatic breast cancer (MEDIOLA): an open-label, multicentre, phase 1/2, basket study. *Lancet Oncol.* **21**, 1155–1164 (2020).
194. Hamanishi, J., Mandai, M., Ikeda, T., Minami, M., Kawaguchi, A., Murayama, T., Kanai, M., Mori, Y., Matsumoto, S., Chikuma, S., Matsumura, N., Abiko, K., Baba, T., Yamaguchi, K., Ueda, A., Hosoe, Y., Morita, S., Yokode, M., Shimizu, A., *et al.* Safety and antitumor activity of Anti-PD-1 antibody, nivolumab, in patients with platinum-resistant ovarian cancer. *J. Clin. Oncol.* **33**, 4015–4022 (2015).
195. Liu, Y. L., Selenica, P., Zhou, Q., Iasonos, A., Callahan, M., Feit, N. Z., Boland, J., Vazquez-Garcia, I., Mandelker, D., Zehir, A., Burger, R. A., Powell, D. J., Friedman, C., Cadoo, K., Grisham, R., Konner, J. A., O'Cearbhaill, R. E., Aghajanian, C., Reis-Filho, J. S., *et al.* BRCA Mutations, Homologous DNA Repair Deficiency, Tumor Mutational Burden, and Response to Immune Checkpoint Inhibition in Recurrent Ovarian Cancer. *JCO Precis. Oncol.* **4**, 665–679 (2020).
196. Vaquero-Garcia, J., Barrera, A., Gazzara, M. R., Gonzalez-Vallinas, J., Lahens, N. F., Hogenesch, J. B., Lynch, K. W. & Barash, Y. A new view of transcriptome complexity and regulation through the lens of local splicing variations. *Elife* **5**, 1–30 (2016).
197. Phillips, J. W., Pan, Y., Tsai, B. L., Xie, Z., Demirdjian, L., Xiao, W., Yang, H. T., Zhang, Y., Lin, C. H., Cheng, D., Hu, Q., Liu, S., Black, D. L., Witte, O. N. & Xing, Y. Pathway-guided analysis identifies Myc-dependent alternative pre-mRNA

splicing in aggressive prostate cancers. *PNAS* **117**, 5269–5279 (2020).

198. Spellman, S., Brewerton, S., Xie, C., Kirkness, E. F., Piper, J., Venter, J. C., Long, T., Turpaz, Y., Yeo, Z. X., Scheuermann, R. H., Biggs, W. H., Vierra-Green, C., Howard, S., Telenti, A., Bloom, K., Wong, M. & Brady, C. Fast and accurate HLA typing from short-read next-generation sequence data with xHLA. *Proc. Natl. Acad. Sci.* **114**, 8059–8064 (2017).
199. Richman, L. P., Vonderheide, R. H. & Rech, A. J. Neoantigen Dissimilarity to the Self-Proteome Predicts Immunogenicity and Response to Immune Checkpoint Blockade. *Cell Syst.* **9**, 375-382.e4 (2019).