FACTORS REGULATING RETROTRANSPOSON EXPRESSION:

Uncovering a Novel BRCA1 Related Mechanism in Ovarian Cancer

Maisa Alkailani

A thesis submitted in partial fulfillment of the requirements for the

Doctorate in Philosophy degree

in

Cellular and Molecular Medicine

Department of Cellular and Molecular Medicine

Faculty of Medicine

University of Ottawa

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADPr: ADP-ribose

AGS: Aicardi-Goutières syndrome

AIM2: Absent in Melanoma 2

ATM: Ataxia-telangiectasia mutated protein

ATR: Ataxia telangiectasia

BARD1: BRCA1-Associated RING Domain-1

BF: Bayes factor

cGAS: cyclic GMP-AMP Synthase

ChIP: Chromatin Immunoprecipitation

Co-IP: Co-Immunoprecipitation

COBRA1: Cofactor of BRCA1

CpG: Cytosine-phosphate-Guanine

CPM: Counts per million mapped reads

DDR: DNA damage response

DNMT: DNA methyltransferase enzymes

DNMTi: DNMT inhibitors

dNTPase: Deoxynucleoside triphosphate triphosphohydrolase

dNTPs: Deoxynucleoside triphosphates

DRIP: DNA-RNA Immunoprecipitation

DRP: Discordant read pair

DSBs: Double strand breaks

dsDNA: Double stranded DNA

dsRNA: Double stranded RNA

EN: Endonuclease

ER: Estrogen receptor

ERE: Estrogen Response Elements

EUL1Db: The European database of L1HS retrotransposon insertions

EVs: extracellular vesicles

FCCP: Carbonyl cyanide 4-[trifluoromethoxy]phenylhydrazone

GDC: Genomic Data Commons

GO: The Gene Ontology

HDACi: histone deacetylation inhibitors

HERVs: Human endogenous retroviruses

HR: Homologous recombination

ICGC: The International Cancer Genome Consortium

IFN: Interferon

KRAB-ZFPs: Krüppel- associated box domain zinc finger proteins

L1: Long Interspersed Element 1

LTR: Long terminal repeats

MAVS: mitochondrial antiviral signaling protein

MDA5: Melanoma Differentiation-Associated protein 5

MEIs: mobile element insertions

MELT: The Mobile Element Locator Tool

MOV10: Moloney leukemia virus 10

NAD+: nicotinamide adenine dinucleotide

NHEJ: Non-homologous end-joining

NNRTIs: Non-nucleoside reverse transcriptase inhibitors

NQO1: quinone reductase (NAD(P)H:(quinone acceptor) oxidoreductase

NRTIs: Nucleoside reverse transcriptase inhibitors

OCR: Oxygen consumption rates

Oct1: Octamer transcription factor 1

ORF1p: L1 ORF1 protein

ORF2p: L1 ORF2 protein

ORFs: Open reading frames

PALB2: partner and localizer of BRCA2

PARP: Poly (ADP-ribose) polymerase 2-5A: 2',5'-oligoadenylate

PARPi: PARP inhibitors

piRNA: Piwi-interacting RNA

polyA: Polyadenosine-rich

PRR: pattern recognition receptors

RIG-I: Retinoic-acid-Inducible Gene I

RNP: Ribonucleoprotein particle

RT: Reverse transcriptase

SAMHD1: SAM domain and HD domain 1

SETX: Senataxin

SGs: stress granules

SINEs: Short Interspersed Elements

SLE: Systemic Lupus Erythematosus

SR: Split read

ssDNA: Single stranded DNA

ST18: Suppression of Tumorigenicity18

StepAIC: Step Akaike Information Criteria)

TCGA: The Cancer Genome Atlas project

TEs: Transposable elements

TLR: Toll-Like Receptor

***TP53*:** The tumor suppressor 53

TPRT: target-primed reverse transcription

TREX1: Three-prime repair exonuclease 1

TSDs: Target site duplications

TSS: Transcription Start Sites

U1 snRNP: U1 small nuclear ribonucleoprotein particle

UTR: Untranslated region

WES: Whole exome sequencing

WGS: Whole genome sequencing

YY1: Yin Yang 1

ZAP: Zinc-finger antiviral protein

ABSTRACT

Retrotransposons constitute about a third of our genome. It is challenging to identify the causes and consequences of retrotransposon expression in human disease due to hundreds of active genomic copies and poor conservation across species. We profiled genomic insertions of retrotransposons in ovarian cancer. RNAs exhibiting Bayesian correlation with retrotransposon RNA were analyzed to identify potential causes and consequences of retrotransposon expression in ovarian and breast cancers. This strategy found divergent inflammatory responses associated with retrotransposon expression in ovarian and breast cancer. It identified new factors inducing the expression of endogenous retrotransposons, including anti-viral responses and the tumor suppressor BRCA1. In cell lines, mouse ovarian epithelial cells and patient-derived tumor spheroids, BRCA1 promoted the accumulation of retrotransposon RNA and facilitated transcription of active families of retrotransposons and their insertion into the genome. Intriguingly, elevated retrotransposon expression predicted survival in ovarian cancer patients. Retrotransposons are part of a complex regulatory network in ovarian cancer, including BRCA1 contributing to patient survival. The above-described analysis strategy could also be used to identify the regulators and impacts of retrotransposons in various contexts of biology and disease in humans.

RÉSUMÉ

Les rétrotransposons constituent environ un tiers de notre génome. Il est difficile d'identifier les causes et les conséquences de l'expression des rétrotransposons dans les maladies humaines en raison de centaines de copies génomiques actives et de leur mauvaise conservation à travers les espèces. Nous avons profilé les insertions génomiques de rétrotransposons dans le cancer de l'ovaire. En outre, pour identifier les causes et les conséquences de l'expression du rétrotransposon dans les cancers de l'ovaire et du sein, nous avons analysé les ARN présentant une corrélation bayésienne avec l'ARN du rétrotransposon. Cette stratégie trouve des réponses inflammatoires divergentes associées à l'expression du rétrotransposon dans le cancer de l'ovaire et du sein et identifie de nouveaux facteurs induisant l'expression de rétrotransposons endogènes, y compris les réponses antivirales et le suppresseur de tumeur BRCA1. Dans les lignées cellulaires, les cellules épithéliales ovariennes de souris et les sphéroïdes tumoraux dérivés du patient, BRCA1 favorise l'accumulation d'ARN rétrotransposon. BRCA1 favorise la transcription des familles actives de rétrotransposons et leur insertion dans le génome. Curieusement, une expression élevée du rétrotransposon prédit la survie chez les patientes atteintes d'un cancer de l'ovaire. Les rétrotransposons font partie d'un réseau de régulation complexe du cancer de l'ovaire, y compris BRCA1, qui contribue à la survie des patientes. En outre, la stratégie d'analyse décrite peut être utilisée pour identifier les régulateurs et les impacts des rétrotransposons dans divers contextes de biologie et de maladie chez l'homme.

PREFACE

This thesis is presented in the traditional format in accordance with the National Library of Canada guidelines. It contains a collaborative work between the author and other researchers contributed to the research project in the published paper (Alkailani et al., 2021) in which the author helped in designing experiments and performed them (unless mentioned otherwise), analyzed experimental data and helped writing the manuscript. Gareth Palidwor performed the bioinformatic analysis of all patient data including methylation, retrotransposon RNA and genomic insertion levels. Ariane Poulin helped performing retrotransposon expression quantification and *Alu* colony formation assays for certain candidate genes. Raghav Mohan and David Pepin prepared patient-derived spheroids used in the study. James Taylor was acknowledged to prepare Northern blot long probes. Barbara Vanderhyden helped in experimental design, data analysis and writing the manuscript. Derrick Gibbings led and conceived the whole project, including experiments design and analysis, performed secondary analysis of correlation data and helped in manuscript writing.

The research work demonstrated in this thesis was generously supported by Qatar National Research Fund (QRLP9-G-3330014 to Maisa Alkailani), US Department of Defense CDMRP Ovarian Cancer Research Program (W81XWH-17-1-0212 to David Pepin), Canadian Cancer Society Research Institute (702978 to Derrick Gibbings), Canadian Institutes of Health Research (142174 to Derrick Gibbings), and Canadian Institutes of Health Research (326557 to Barbara Vanderhyden).

Research work presented in this thesis was performed in the Department of Cellular and Molecular Medicine, Faculty of Medicine and Ottawa Institute for Systems Biology, University of Ottawa, Ottawa, Ontario; Bioinformatics and Cancer Therapeutics Program Ottawa Hospital Research Institute, Ottawa, Ontario; Pediatrics Surgical Research Laboratories, Massachusetts General Hospital and Department of Surgery, Harvard Medical School, Boston, MA.

ACKNOWLEDGEMENTS

Nothing becomes impossible when having great people around you, share the pains before taking the gains. My gratitude is to everyone who supported me and contributed to this work until it has been accomplished and become a reality.

To whom financially secured and sponsored my studies, great thanks to my program in Qatar (**Qatar Research Leadership Program (QRLP)**) for giving me the opening to learn, progress, and excel in my research capabilities. A special thank you goes to QRLP unique team who were working behind the scenes to make my life easier during the long journey: **Dr. Ayman Bassil**, **Ms. Noha Al-Okka**, and **Ms. Maria Susi Estacio,** all your efforts are highly valued, and your hard work is paying off delightful fruits.

Special and sincere thanks to my supervisor **Dr. Derrick Gibbings**, who gave me a nice niche to grow in and helped to shape a better version of myself. He is whom you call a bright scientist, a great mentor and a good leader. I found his support throughout stressful times, when I needed technical help, he was at the bench to learn from, and his excellent coaching was endless. He formulated my passion for the retrotransposons field by his guidance and brilliant ideas that were most fruitful.

My warm thanks to all my friends in **The Gibbings Lab**, including the lab alumni, for never failing to help and advice whenever I needed it, for their support in difficult times

and for the interesting conversations we had together. I had my unforgettable best days in this lab among these outstanding people.

Great thanks to **Dr. Barbara Vanderhyden**, my co-supervisor, who left a unique mark in my learning journey with her science passion, kindness, caring and support. For allowing me to join their meetings and communicate my research within her team. I want to extend my thanks to all of her lab members for their constructive feedback on my project.

I want to thank all of my Thesis Advising Committee (TAC) members, namely, **Dr. Tommy Alain**, **Dr. Jeffrey Dilworth**, and **Dr. Carolina Ilkow**, who contributed to improving this work by giving me a hard time during TAC meetings for their kind advice and valued guidance.

My tender thanks to my parents, **Ismael Alkailani** and **Rasmiya Al-Aidi**, my life motivators and great influencers for being my best role models of working hard and trying their best to make a big difference every day. They showed me practically how to succeed in reaching goals. I learnt from them that I could achieve success by having the excellent character of being patient and persistent. Success is a journey filled with varied experiences that shape the future and begins with one's pursuit.

Words cannot describe my heartfelt thanks to my partner in life and my best friend, **Abdurrahman Alaydi**, for being at my side throughout my studies, covering my back, caring, supporting, and inspiring. His perseverance and resilience to overcome career

challenges encouraged me to do similar. I learnt from him to think about every new day as a unique opportunity to start again and continue the adventure.

My warm thanks are to our five children (**Albaraa**, **Jomana**, **Sara**, **Maria**, and **Jasmine**) for beautifying my days, boosting my energy every day and for their friendly cooperation and pleasant understanding till the end, especially during the COVID-19 time. My productive days started with your hugs and smiles. I'm confident that one day you will appreciate the value behind your mother's struggles in grad school as much as I'm convinced that you will succeed in reaching your big goals.

Last but not least, a special thank you note to all my **family members**, **relatives**, and **friends** who supported me throughout the journey and had significant impacts on completing this work in different ways.

I'm so grateful to be surrounded by unique people; every one of them played an essential role in helping, supporting and motivating me to complete what I decided to start. I could not achieve anything of this without everyone around me.

COPYRIGHTED CONTENTS

Thesis's text, figures, and tables contain parts written and presented *de novo*, and others reproduced with permission. Chapter 1, background and discussion sections of Chapters 2-5, section 4.2.2 and Chapters 6-8 were written and created *de novo* unless otherwise is indicated.

Copyright licenses (5002611244482, 5002620366737, 5002621161678) were obtained from Oxford University Press- NAR Cancer to re-use the following with citation: figures, supplementary figures, figure captions, supplementary tables, and sections of materials & methods from the published paper, Alkailani, M. et al., 2021, A genome-wide strategy to identify causes and consequences of retrotransposon expression finds activation by BRCA1 in ovarian cancer, NAR Cancer.

Copyright licenses (5073120506815, 5071521089348) were obtained from Springer Nature to reproduce parts of published figures. Copies of the mentioned above licenses are attached to Appendix F.

1 CHAPTER 1: GENERAL INTRODUCTION

1.1 Transposable Elements Definition and Classification

Transposable elements (TEs) are mobile DNA sequences that can move from one genomic location to another in a process referred to as "transposition" (Bourque et al., 2018). Transposition in the genome is facilitated by one or more proteins encoded by a TE (Kazazian Jr. and Moran, 2017). The mechanism by which a transposition occurs is defined separately for each TE class. As represented in Figure 1-1 below, TEs are categorized into two broad classes: DNA transposons and retrotransposons, based on their transposition intermediate and mobility mechanism (Kazazian Jr. and Moran, 2017). DNA transposons are DNA sequences that use element-encoded transposases to move from one genomic location to another by a cut-and-paste mechanism (Kazazian Jr. and Moran, 2017). A retrotransposon element moves and inserts into a new genomic location by a copy-and-paste mechanism using an RNA intermediate (Boeke et al., 1985; Kazazian Jr. and Moran, 2017).

1.1.1 DNA Transposons

According to early reports of human genome sequencing, DNA transposons make up about 3% of our genome (Lander et al., 2001). There is no evidence of DNA transposon activity generating new insertions in the human genome in the last 37 million years (Pace and Feschotte, 2007). Therefore, in this study, the focus will be on retrotransposons, the active transposing class of TEs that constitute ~45% of the human genome (Lander et al., 2001).

1.1.2 Retrotransposons

DNA of a retrotransposon element is transcribed to RNA. In the "retro" step, this RNA is copied back into DNA by a reverse-transcriptase (RT) enzyme encoded by the retrotransposon. Then this reverse-transcribed DNA integrates into a new genomic location (Kazazian Jr. and Moran, 2017). According to Long Terminal Repeats (LTR), presence or absence in their sequences, retrotransposons are subdivided into two subtypes: LTR and non-LTR elements (Kazazian Jr. and Moran, 2017).

1.1.2.1 LTR Elements

Retrotransposons containing LTRs, also known as human endogenous retroviruses (HERVs), are autonomous protein-encoding elements that account for ~ 8% of the human genome (Kazazian Jr. and Moran, 2017). Their genetic structure is comparable to exogenous retroviruses; each HERV element consists of two identical LTRs, flanking open reading frames (ORFs), which encode proteins essential for viral replication (Kazazian Jr. and Moran, 2017). HERV contains gag, pol, and env genes to encode group-specific retroviral antigen (Gag), reverse transcriptase (Pol) and envelope protein (Env), respectively (Kazazian Jr. and Moran, 2017). Most HERV elements are non-functional due to accumulated mutations or internal recombination, resulting in a solitary LTR (Jern and Coffin, 2008). However, evidence suggests the recent insertion activity of HERV elements within the human population in polymorphic loci (Wildschutte et al., 2016).

1.1.2.2 Non-LTR Elements

Retrotransposons lacking LTR include Long Interspersed Elements (LINEs) and Short Interspersed Elements (SINEs). Of these, the most active elements retrotransposing in the human genome include autonomous LINE-1 (L1) from LINEs and non-autonomous *Alu* from SINEs (Kazazian Jr. and Moran, 2017).

1.1.2.2.1 LINE-1

More than 30% of the human genome is derived directly or indirectly from L1 retrotransposon activity (Hancks and Kazazian, 2012). Due to point mutations, rearrangements, or truncations, of the 500,000 L1 copies present in the human genome, only a subset of 80-100 elements is currently active (Hancks and Kazazian, 2012). The full-length human L1 element is 6 kb having bidirectional promoter activities. The 5' untranslated region (UTR) of L1 contains two promoters, one in the sense direction and the other in the antisense orientation (Hancks and Kazazian, 2012; Burns, 2017). L1 sense promoter transcribes the two ORFs (ORF1 and ORF2) (Hancks and Kazazian, 2012; Burns, 2017). The antisense promoter transcribes a primate-specific ORF (ORF0) in the opposite orientation of L1 (Denli et al., 2015). L1 contains short inter-ORF spacers between ORFs and ends with a 3' UTR polyadenosine-rich (polyA) tract (Hancks and Kazazian, 2012; Burns, 2017). ORF0 was shown to enhance L1 mobility and contribute to L1-mediated diversity by producing chimeric transcripts and fusion proteins depending on proximal genes (Denli et al., 2015). ORF1 encodes a 40-kDa protein (ORF1p) with nucleic acid chaperone and RNA binding activities (Martin et al., 2003). ORF2 encodes a 150-kDa protein (ORF2p) that has endonuclease (EN) and RT activities (Mathias et al.,

1991; Feng et al., 1996). ORF1p, ORF2p and the 3' polyA tract were shown to be required for efficient retrotransposition of L1 in *cis* and *Alu* elements in *trans* (Dewannieux et al., 2003; Doucet et al., 2015).

Transcription of L1 mRNA from a genomic copy is a start point of its life cycle (Kazazian Jr. and Moran, 2017). RNA is exported to the cytoplasm where both ORF1p and ORF2p proteins are translated (McMillan and Singer, 1993). Then they bind the retrotransposon RNA to form a ribonucleoprotein particle (RNP), which is the core of the retrotransposition machinery (Kulpa and Moran, 2006; Doucet et al., 2010). After being imported into the nucleus, this RNA facilitates L1/ *Alu* retrotransposition via two distinct pathways (Viollet et al., 2014). The canonical pathway is called target-primed reverse transcription (TPRT) (Luan et al., 1993; Cost et al., 2002) in which the L1 EN activity produces a nick at a recognized target site in the genomic DNA (Feng et al., 1996; Cost and Boeke, 1998). It preferentially cuts DNA at a consensus sequence 5'-TTTT/A-3' or variants (Cost and Boeke, 1998). Then, using the retrotransposon RNA as a template, L1 RT moiety extends the unbound 3'-OH group from DNA to begin reverse transcription starting within the polyA tail of the retrotransposon RNA (Cost and Boeke, 1998; Kazazian Jr. and Moran, 2017). Retrotransposition can also occur via an endonuclease-independent pathway or non-classical L1 insertion (Morrish et al., 2002; Sen et al., 2007). In this pathway, endonuclease cleavage is not required, and the reverse transcription is initiated at pre-existing DNA break regions (Morrish et al., 2002; Sen et al., 2007). Retrotransposition insertions at telomeric regions are examples of this pathway (Kopera et al., 2011). Research is still ongoing to discover the details of the following steps of both

pathways, such as second-stranded DNA synthesis and ligation of the 3' ends to the target DNA (Viollet et al., 2014; Kazazian Jr. and Moran, 2017).

1.1.2.2.2 Alu

Alu elements are primate-specific retrotransposons that comprise ~ 11% of the human genome (Deininger, 2011). The most recent *Alu* amplification in primate lineages has been attributed to a series of *Y* subfamilies, with *Ya5* and *Yb8* dominating humans (Deininger, 2011). Each *Alu* element is approximately 300 nt long composed of two dimers ancestrally derived from the 7SL RNA (that generates RNA of the signal recognition particle) separated by a short polyA sequence, and its 3' end is occupied by a longer polyA tail (Deininger, 2011). *Alu* element transcription starts from an internal RNA polymerase III promoter located at its 5' end (Deininger, 2011). Because the *Alu* element has no transcription terminator, a TTTT terminator sequence in a nearby genomic location is used to terminate its transcription process (Deininger, 2011). *Alu* elements do not encode proteins but rather hijack L1 elements to mediate their retrotransposition (Deininger, 2011). *Alu* utilizes the retrotransposition machinery slightly different from L1 since it depends mostly on ORF2p (Dewannieux et al., 2003; Deininger, 2011). *Alu* can contribute to genomic instability and disease by inserting in introns, 3'UTRs of genes, or intergenic regions (Deininger, 2011).

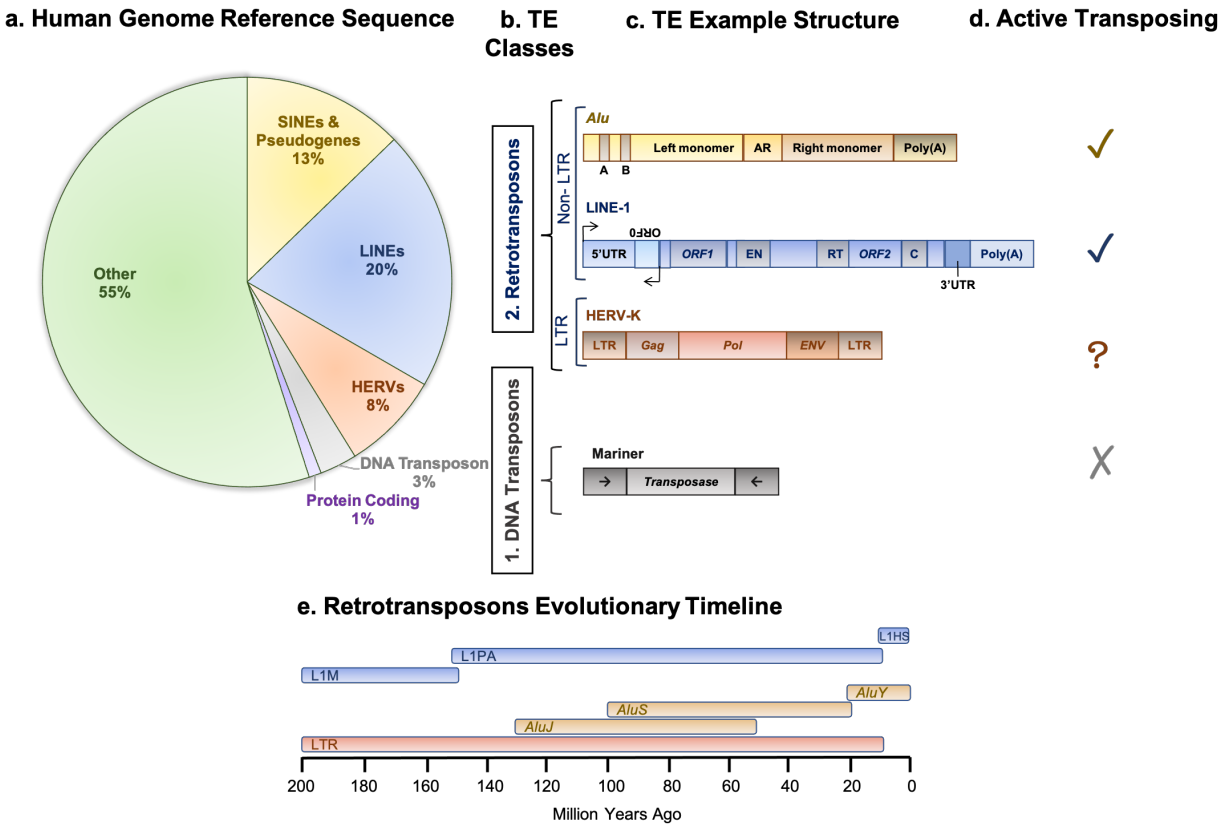


Figure 1-1Transposable Elements Classes, Structure and Activity.
a. A pie chart represents fractions of human genome reference sequence constituents as described in (Lander et al., 2001). **b**. TE main categories: first class, DNA transposons and second class, retrotransposons. The latter is subcategorized into elements having or lacking LTRs. LTR containing elements include HERVs family and non-LTR elements include SINEs, LINEs and pseudogenes families. **c.** Outline diagrams to represent structure of example elements per classes of transposable elements, Alu element from SINEs family is composed of two monomers separated by adenosine-rich (AR) linker. The left monomer contains an internal RNA polymerase III promoter (bars labeled A and B), and the right monomer is followed by a poly(A) tail. L1 element is a protein coding element of the LINEs family, has an internal promoter in its 5' untranslated region (5'UTR) followed by primate specific antisense region (ORF0), and regions encoding L1 proteins (ORF1 and ORF2). ORF1p is a nuclear binding protein and ORF2p has EN, RT and cysteine-rich (C) domains. L1 element is ended by a poly(A) tail in its 3' untranslated region (3'UTR). HERV-K element of the HERVs family, contains two LTR regions separated by gag, pol, and env regions. Mariner of the DNA transposons class encodes transposase, an enzyme that binds and cut near inverted repeats flanking the element (denoted by little arrows). **d.** Symbols representing recent transposing activity in modern human as reported in literature. ✓ sign denotes existing evidence of current activity, ? sign denotes no certain activity and X sign denotes no reported evidence. **e.** Approximate time spans of recent retrotransposons activity. L1HS, L1PA, L1M are subfamilies of L1, *AluY*, *AluS*, and *AluJ* are subfamilies of *Alu*. Panels **a, c** and **d** were reproduced with permission from (Kazazian Jr. and Moran, 2017), Copyright Massachusetts Medical Society. Panel **e** was reproduced with permission from (Burns, 2017), Copyright Springer Nature.

1.2 A Flashback on Transposable Elements Discovery

"You just know sooner or later, it will come out in the wash, but you may have to wait some time." This statement was made by Dr. Barbara McClintock upon receiving the

Nobel Prize to recognize her discovery of transposable elements (Ravindran, 2012). In the 1900s, researchers re-discovered Mendel's work on inheritance and started to study genetics and inheritance using maize and fruit fly models in addition to pea plants (Ravindran, 2012). During that era, the widely accepted concept about genes was that they were static units of heritability arranged linearly on a chromosome as beads on a string (Ravindran, 2012). After studying an undergraduate degree in agriculture and a graduate degree in botany, Dr. McClintock developed and refined staining techniques to visualize the ten maize chromosomes during her work as an instructor at Cornell's College (Ravindran, 2012). Those techniques paved the way for her discovery of "jumping genes" in the 1940s while working as a full-time researcher at the Carnegie Institution in Cold Spring Harbor (Ravindran, 2012). It is noteworthy that Dr. McClintock had to wait thirty-five years until the scientific community acknowledged her significant contribution (Ravindran, 2012).

She revealed a particular chromosome breakage event that occurred recurrently at the same locus on the maize ninth chromosome (McClintock, 1950). She called it: "dissociation" or "*Ds*" locus, which could change position within the chromosome (McClintock, 1950). Moreover, the chromosome breakage at the *Ds* locus required another dominant locus that could initiate its transposition, too (McClintock, 1950). Which she named locus "*Activator,"* or "*Ac,"* and she discovered that an *Ac* element could activate *Ds* chromosome breakage at a different location or on a different chromosome (McClintock, 1950).

Other pioneering scientists built on McClintock's research in TEs, notably Roy Britten and Eric Davidson, who speculated about TE's function (Britten and Davidson, 1969). They thought that TEs play a role in the regulation of gene expression and promote the generation of different cell types and biological structures based on the genomic site they insert themselves in (Britten and Davidson, 1969; Chuong et al., 2017).

These early observations and speculations of McClintock, Britten and Davidson were primarily ignored within the scientific community, and TEs were considered a part of "junk" DNA for decades (Chuong et al., 2017). The interest in studying these elements was weak until the completion of the human genome project, which revealed surprising discoveries (Lander et al., 2001). TEs and sequences derived from them comprise about one-half of the human genome, while no more than 1% of it is composed of protein-coding exons (Lander et al., 2001). Since then, this "junk" material composed of TEs has attracted significant attention from researchers to explore their biology and impacts.

1.3 Transposable Elements and Structural Variation

Any two human individuals are ~ 99.9% identical in their DNA sequence, as revealed by the human genome analysis (Feuk et al., 2006). Studying a small fraction of the genome that constitutes the genetic variation between individuals became feasible with the recent advances in experimental and computational technologies (Feuk et al., 2006). Structural variants can have balanced or unbalanced forms that involve no net gain/loss in the genetic material or lead to copy number variations, respectively (Feuk et al., 2006). Inversions and translocations are examples of the balanced variants, whereas deletions, duplications,

and insertions are examples of the unbalanced ones (Feuk et al., 2006). Structural variants can arise as consequences of multiple molecular mechanisms that include DNA replication, recombination, and repair-associated processes as well as mobile element insertions (Weischenfeldt et al., 2013).

1.3.1 DNA Damage and Retrotransposons

DNA damage is a primary endogenous source of structural variation (Tubbs and Nussenzweig, 2017). Each human cell is subject to ~ 70,000 DNA lesions per day (Tubbs and Nussenzweig, 2017). DNA damage response (DDR) has evolved in mammalian cells to detect DNA lesions and mediate their repair through different pathways depending on the type of DNA break (Jackson and Bartek, 2009).

1.3.1.1 DNA Damage Repair General Pathways

Serial catalytic events mediated by several involved proteins can repair most lesions (Jackson and Bartek, 2009). Two key DDR- signaling components are the protein kinases Ataxia-telangiectasia mutated protein (ATM) and Ataxia telangiectasia (ATR) (Jackson and Bartek, 2009). Double-strand breaks (DSBs) recruit and activate ATM, whereas replication protein A (RPA)-coated single-stranded DNA (ssDNA) recruits and activates ATR (Jackson and Bartek, 2009). Mismatches and insertion/deletion loops trigger single-strand incisions. The mismatch repair machinery repairs these incisions using nuclease, polymerase and ligase enzymes(Jackson and Bartek, 2009). The nucleotide excision repair (NER) system recognizes helix-distorting base lesions (Jackson and Bartek, 2009). The damage is excised as a 22–30 base oligonucleotide, generating ssDNA, repaired by DNA polymerases and associated factors and terminated by ligation (Jackson and Bartek, 2009).

Two significant mechanisms can repair DNA DSBs: non-homologous end-joining (NHEJ) and homologous recombination (HR). NHEJ is an error-prone mechanism that operates in any phase of the cell cycle. Ku proteins (including XRCC6) operate to recognize DSBs then bind and activate DNA protein kinases (PKcs). They recruit and activate polymerases and DNA ligase IV to complete the process. On the other hand, HR is usually restricted to the S and G2 phases of the cell cycle because it uses sister-chromatid sequences as templates to mediate accurate repair mechanisms (Jackson and Bartek, 2009).

HR, in summary, is initiated by an ssDNA generation that is promoted by components of the MRN complex (MRE11– RAD50– NBS1). In a subsequent event catalyzed by RAD51, RAD51 related proteins (including DMC1) and the breast cancer susceptibility proteins (BRCA1 and BRCA2), the ssDNA invades an undamaged DNA template and pairs with a homologous region to initiate HR (Jackson and Bartek, 2009). Then polymerases, nucleases, helicases and other components take action (Jackson and Bartek, 2009). Finally, DNA ligation and substrate resolution occur (Jackson and Bartek, 2009).

1.3.1.2 BRCA1 structure and function

BRCA1 acts as a caretaker tumor suppressor that maps to the human chromosome 17q 21.3 (Hall et al., 1990). Disruption of both copies of the *BRCA1* gene in the mouse germline impedes cell proliferation and leads to early embryonic lethality (Hakem et al., 1996). 72% and 44% of *BRCA1* mutation carriers develop breast and ovarian cancer by

the age of 80 years, respectively (Kuchenbaecker et al., 2017). *BRCA1* encodes a large protein of 1863 amino acids that physically interacts with numerous partner proteins directly or indirectly (Huen et al., 2010). BRCA1 protein is regulated by nuclear-cytoplasmic shuttling, with most of its described functions being predominantly nuclear (Ruffner and Verma, 1997; Henderson, 2005). BRCA1 is directly involved in the HR-mediated DNA repair pathway of DSBs (Scully et al., 1997a).

BRCA1 protein contains multiple functional domains, a conserved amino-terminal RING domain, a coiled-coil (CC) domain and carboxy-terminal tandem BRCT domains (Mark et al., 2005). BRCA1 forms a heterodimer with BARD1 (BRCA1-Associated RING Domain-1) by an interaction between the RING domains of both proteins (Mallery et al., 2002). BRCA1/BARD1 complex acts as an E3 ubiquitin ligase (Mallery et al., 2002). Evidence suggests that mutations in the RING domain of BRCA1 can contribute to tumorigenesis by abrogating its E3 ubiquitin ligase activity (Hashizume et al., 2001; Ruffner et al., 2001).

In response to a DSB, BRCA1, via its CC domain, interacts with PALB2 (partner and localizer of BRCA2) protein that recruits BRCA2 to the lesion to complete the repair process (Zhang et al., 2009). The tandem BRCT domains recognize a phosphorylated serine in a pSXXF motif which exists in three proteins (Abraxas, CtBP-interacting protein (CtIP) and BRCA1 Interacting Helicase 1 (BRIP1)) (Prakash et al., 2015). Abraxas and CtIP proteins form independent complexes with BRCA1 which are essential for different stages of HR-mediated repair (Wang et al., 2007; Li and Yu, 2013). As represented in

Figure 1-2, HR repair of DSBs starts by recruitment of BRCA1 to the DNA lesion site with the help of Abraxas. BRCA1 interacts with BARD1, CtIP and MRN complex to prime DNA resection and generate ssDNA (Moynahan et al., 1999; Takaoka and Miki, 2018). BRCA1-PALB2- BRCA2 complex then loads RAD51 at this ssDNA, where HR continues as described earlier (Takaoka and Miki, 2018).

BRCA1 is involved in numerous other roles in the cell, including transcription regulation (Prakash et al., 2015). Evidence showed that BRCA1 could bind RNA Polymerase II and activate transcription (Scully et al., 1997b; Anderson et al., 1998). BRCA1 associates with the RNA polymerase II holoenzyme transcription complex (holo-pol) via its BARD1 interaction (Chiba and Parvin, 2002). Although both BRCA1 termini could bind the holo-pol complex, the N-terminus was shown to be more essential for this association than its C-terminus (Chiba and Parvin, 2002). Being a component in the holo-pol complex, BRCA1 acts as a transcriptional coactivator of multiple factors such as p53 (Zhang et al., 1998), NF-κB (Benezra et al., 2003), and STAT1 (Ouchi et al., 2000). BRCA1 can also transcriptionally repress the activity of other factors such as c-Myc (Wang et al., 1998) and estrogen receptor (ER) (Zheng et al., 2001). BRCA1 preferentially associates with transcription start sites (TSS) in the human genome consistent with its functional roles in regulating transcription (Gardini et al., 2014). BRCA1 was shown to stimulate transcription without the requirement for a DNA-tethering function (Nadeau et al., 2000). Therefore, BRCA1 regulation of transcription is suggested to occur through a mechanism alternative to RNA polymerase II recruitment by modulating its phosphorylation (Moisan et al., 2004). BRCA1 was shown to negatively modulate the phosphorylation levels of

RNA polymerase II carboxy-terminal domain *in vitro* (Moisan et al., 2004). When hypo-phosphorylated, RNA polymerase II is recruited to a target promoter to form a stable preinitiation complex (Komarnitsky et al., 2000). On the other hand, when RNA polymerase II is hyper-phosphorylated, it associates with the gene coding region (Komarnitsky et al., 2000). BRCA1 preferentially associates to the hyper-phosphorylated RNA polymerase II (Krum et al., 2003). Another mechanism could include inhibiting the coupled transcription-RNA processing machinery and facilitating associated DNA damage repair (Kleiman et al., 2005). This mechanism involves targeting the elongated RNA polymerase II for ubiquitination by the BRCA1/BARD1 heterodimer (Kleiman et al., 2005). BRCA1 was also shown to interact with a cofactor of BRCA1 (COBRA1) that constitutes the B subunit in the negative elongation factor (NELF) complex (Nair et al., 2016). This complex acts as a regulator of transcription elongation that pauses RNA polymerase II (Kwak and Lis, 2013), leading to either decreased or increased transcription (Sun et al., 2011). The BRCA1-COBRA1 interaction is independent of DNA damage repair pathways and significant for normal mammary gland development in mice (Nair et al., 2016).

In response to transcription-associated genetic instability, BRCA1 acts to eliminate transcriptional intermediates (Gardini et al., 2014; Hatchi et al., 2015). BRCA1 regulates the transcription of actively transcribed genes by binding R-loops (RNA-DNA hybrids) at promotors and transcription pause sites (Gardini et al., 2014; Hatchi et al., 2015). Although R-loops are naturally occurring and their formation is necessary for multiple cellular processes, they can compromise the genomic integrity if left unresolved (Hamperl

and Cimprich, 2014; Skourti-Stathaki and Proudfoot, 2014). BRCA1 interacts with senataxin (SETX) to form a complex recruited to R-loop-associated transcription pause sites (Hatchi et al., 2015). The BRCA1/SETX complex is necessary to suppress DNA damage occurring at R-loop sites (Hatchi et al., 2015). BRCA1 mutation leads to R-loop accumulation associated with RNA polymerase II pausing, which was alleviated by COBRA1 depletion (Zhang et al., 2017). The roles of BRCA1 in R-loop resolution and transcription regulation limit tumorigenesis and contribute to its primary function as a tumor suppressor (Hatchi et al., 2015; Zhang et al., 2017).

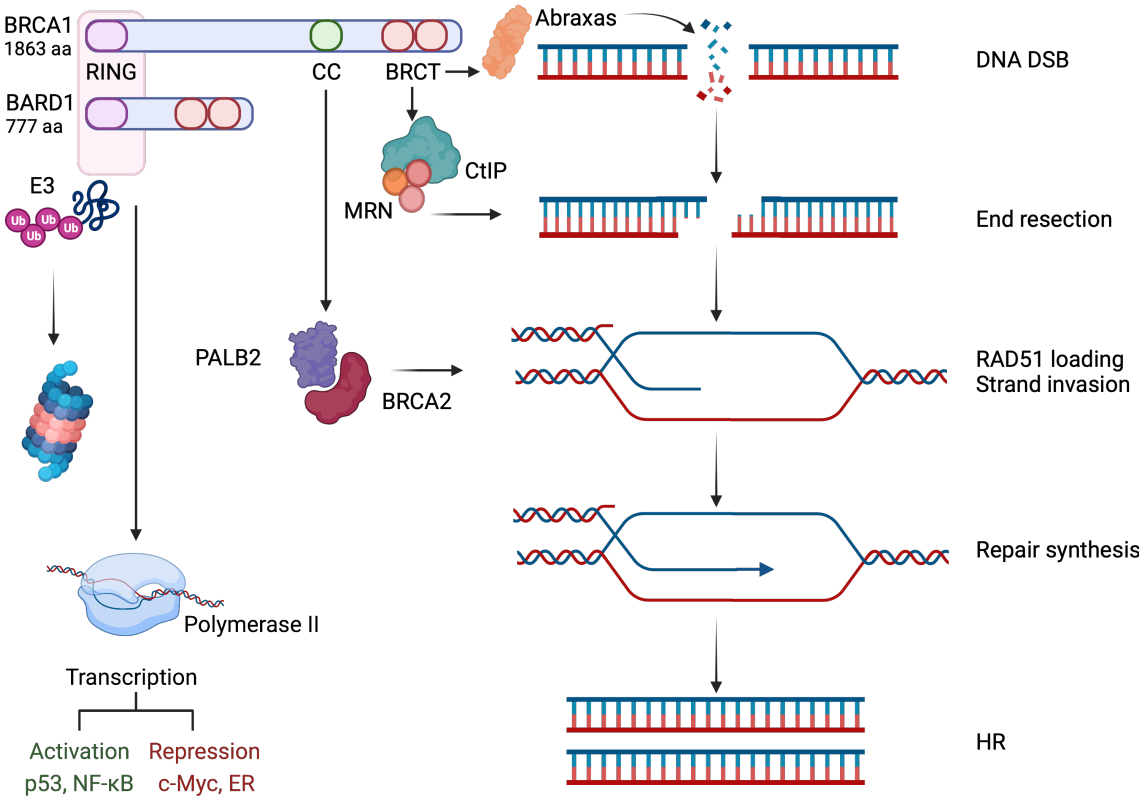


Figure 1-2 BRCA1 structure and role in HR DSB repair pathway
BRCA1 has E3 ubiquitin ligase activity when it interacts with BARD1 via their RING domains. BRCA1 is localized to DNA DSB lesion by an interaction between its C-terminal BRCT repeats and Abraxas protein. DNA end resection occurs by an interaction of BRCA1 with CtIP protein and MRN complex. BRCA1, through its CC domain, interacts with PALB2 that bridges BRCA2 to recruit RAD51 to the formed ssDNA. RAD51 forms a nucleoprotein filament with ssDNA that invades a homologous template DNA. DNA synthesis occurs using the sister chromatid, and the repair

continues to restore the original sequence before the damage. Illustrations were created with BioRender. HR steps were reproduced with permission from (Kaniecki et al., 2018), Copyright Springer Nature.

1.3.1.3 DNA Damage Repair and L1

TEs account for approximately 25% of genetic differences between individual humans (Sudmant et al., 2015), when we account for alterations in the target sequence where TEs insert, frequent DNA breakage at their insertion sites, or their possible unsuccessful insertions (Hastings et al., 2009; Mita and Boeke, 2016). Insertion sites are hot regions for non-homologous recombination and persistent single-strandedness due to extensive transcription, secondary structures and replication pausing (Hastings et al., 2009). L1 EN cleavage of target DNA or other TPRT intermediates can lead to DSBs and activation of consequent DNA repair pathways (Pizarro and Cristofari, 2016). Conversely, the resolution of L1 integration likely needs these signaling pathways (Rodić and Burns, 2013) through L1 second-strand DNA synthesis or DNA ligation (Pizarro and Cristofari, 2016). L1 reporters were shown to induce DSBs, and γ-H2AX foci formation when transiently transfected in cell models and their insertion into new genomic sites is presumed to involve DNA repair mechanisms (Babushok and Kazazian, 2007). Although much research is ongoing in this field, the specific DNA repair pathway required for L1 retrotransposition is unresolved. There are some conflicting data reported in this regard, possibly related to the tools and cell models used to study the DNA repair pathways. Cultured cell models derived from cancers have frequent alterations in DNA repair pathways that could affect the DNA repair mechanisms retrotransposons use during their insertion. In addition, employing vector systems to express TEs with reporter cassettes and exogenous promoters may introduce artificial levels of expression that can modulate the activated pathways (Rodova et al., 2013).

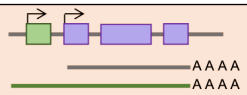
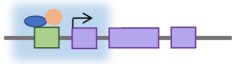
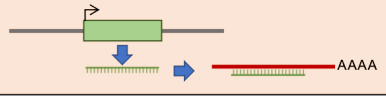
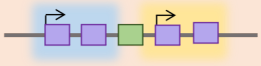
Accumulation of γ-H2AX foci after L1 expression in cells indicated that DSBs were being formed and required functional ATM in the repair process to achieve complete integration of L1 (Gasior et al., 2006). In these experiments, the number of L1- induced DSBs exceeded the predicted number of L1 insertions that occurred, indicating a degree of inefficiency in L1 integration (Gasior et al., 2006). In contrast, evidence from ataxia-telangiectasia (ATM-deficient) patients and cell models showed increased copy numbers of L1 DNA and L1 retrotranspositions, respectively (Coufal et al., 2011). This evidence suggests that ATM is constricting L1 rather than supporting its activity (Coufal et al., 2011). Members of the NHEJ pathway were required for efficient L1 (zebrafish ZfL2-2 and human L1) insertions in chicken cell lines (Suzuki et al., 2009). In comparison, members of the HR and NER pathways were shown to limit L1 retrotransposition in cell models (Brégnard et al., 2016; Servant et al., 2017). L1 insertions were also demonstrated to provide templates for homology-dependent repair of induced DNA DSBs by analyzing whole-genome sequencing of myeloma cell lines, which might contribute to the general structural variation of the human genome (Onozawa et al., 2014).

1.3.1.4 Impact of Retrotransposons on genome structure

De novo retrotransposon insertions can occur in exons, introns or regulatory regions of the genome, disrupting their function, providing new promoter and enhancer regions and contributing to disease (Han et al., 2004; Faulkner et al., 2009). These insertions can exert deleterious, "disruptive," or beneficial "exaptation" effects on the host (Mita and Boeke, 2016). Retrotransposition in intron regions can affect the splicing process by different

mechanisms (Lev-Maor, 2003; Sela et al., 2007). It can provide alternative (donor or acceptor) splice sites, cause exonization (a process by which genes acquire new exons from intronic DNA sequences), or promote exon skipping (Lev-Maor, 2003; Sela et al., 2007). Alternative splicing and exon-acquisition events of the *CHRM3* gene, a member of the muscarinic acetylcholine receptor family, are examples of TE integration into the host genome naturally selected and conserved over generations (Huh et al., 2009). About 62% of exonizations in the human genome are *Alu* derived (Zhang and Chasin, 2006). The insertion of *Alu* into one of *Factor VIII* gene introns resulted in exon skipping and consequent onset of hemophilia A (Ganguly et al., 2003). Table 1-1 outlines mechanisms by which retrotransposons can impact genomic structure and function.

Table 1-1 Mechanisms by which retrotransposons can affect the genome structure.
Retrotransposons can affect the genome structure and impact genomic function by various mechanisms, including providing new promoters, enhancers, or silencer regions, causing alterations in exons, splicing sites or polyA tails, triggering insulation effects, and producing regulatory RNA sequences or new protein.

Retrotransposons Regulatory Effect	Reference	Schematic Illustration
Alternative promoter	(Faulkner et al., 2009) (Conley et al., 2008) (Emera et al., 2012)	
New enhancer/ silencer	(Bejerano et al., 2006) (Franchini et al., 2011)	
Exon disruption/ addition	(Kazazian et al., 1988) (Maksakova et al., 2006) (Shen et al., 2011)	
Alterative polyA	(Lee et al., 2008) (Tang et al., 2000)	
Regulatory RNA production	(Borchert et al., 2006) (Piriyapongsa et al., 2007) (Kapusta et al., 2013)	
New protein production	(Dupressoir et al., 2012)	
Alteration in splicing	(Medstrand et al., 2002) (Lev-Maor et al., 2003) (Sela et al., 2007)	
Deletion/ duplication	(Boissinot et al., 2006) (Song and Boissinot, 2007)	
Insulation	(Wang et al., 2012) (Wang et al., 2015)	

References included: (Faulkner et al., 2009), (Conley et al., 2008), (Emera et al., 2012), (Bejerano et al., 2006), (Franchini et al., 2011), (Kazazian et al., 1988), (Maksakova et al., 2006), (Shen et al., 2011), (Tang et al., 2000), (Borchert et al., 2006), (Piriyapongsa et al., 2007), (Medstrand et al., 2002), (Boissinot et al., 2006), (Song and Boissinot, 2007), (Dupressoir et al., 2012), (Rebollo et al., 2012), (Wang et al., 2012), (Kapusta et al., 2013), (Lee et al., 2008), (Lev-Maor et al., 2003), (Sela et al., 2007), (Wang et al., 2015). Illustrations were reproduced with permission from (Ecco et al., 2017), Copyright Company of Biologists.

1.4 Mechanisms of Retrotransposons Regulation

Previous reports have shown retrotransposon expression, and activity occurs primarily in cells associated with the germline, with little (if any) expression in somatic tissues under

physiological conditions (Belancio et al., 2010). Long considered to be genomic threats to somatic cellular functions, retrotransposons are under control mechanisms that restrict their activity (Van Meter et al., 2014). These regulation mechanisms sometimes fail in cases of age or disease (Van Meter et al., 2014). Advances in molecular techniques, including cell culture retrotransposition assays, made it possible to study factors regulating (restricting or activating) retrotransposons. This progress has significantly increased our understanding of cellular mechanisms controlling retrotransposons activity. Factors restricting retrotransposons fall under one of two categories: cytoplasmic or nuclear. Most factors acting in the cytoplasm are limiting retrotransposons expression by post-transcriptional mechanisms (Goodier, 2016). In contrast, the nuclear factors either suppress the retrotransposons transcription or interfere with their genomic integration (Goodier, 2016) (see Figure 1-3). The following sections will discuss examples of these regulators and their suggested mechanism of retrotransposons suppression.

1.4.1 Transcription factors as regulators of retrotransposons

Several transcription factors were demonstrated to regulate the transcription of retrotransposons by binding their promoters. Examples of these factors include YY1, RUNX3, p53, Oct4, Sox2, Nanog, KLF4, MYC and CTCF (Becker et al., 1993; Yang et al., 2003; Athanikar et al., 2004; Harris et al., 2009; Kunarso et al., 2010; Wang et al., 2014; Grow et al., 2015; Wylie et al., 2016; Sun et al., 2018).

The transcription factor YY1 (Yin Yang 1) has specific binding sites in L1 and can regulate the transcription of L1 (Becker et al., 1993). Another study showed later that

these sites are located in the L1 5'UTR and required for L1 transcription initiation (Athanikar et al., 2004).

Exogenous expression of Runt-domain transcription factor 3 (RUNX3), but not other family members (RUNX1 and RUNX2), increased L1HS transcription and retrotransposition rates in cells (Yang et al., 2003). Site-directed mutation studies revealed that RUNX3 positively regulates the promoter activity of L1HS 5'UTR in both sense and antisense directions (Yang et al., 2003).

The tumor suppressor *TP53*, whose protein (p53) acts as a transcription factor, has many DNA binding sites detected in the promoter region of L1 (Harris et al., 2009). The binding of p53 to L1 through these functional binding sites increased L1 mRNA expression in HCT116 cells (Harris et al., 2009). Conversely, another study that expressed wild-type human p53 in the Drosophila model showed p53 restraining rather than activating L1 transcription (Wylie et al., 2016). This observation was not noticed when common cancer variants of p53 were used (Wylie et al., 2016). Nevertheless, evidence suggests that p53 regulates L1 by binding its promoter (Harris et al., 2009; Wylie et al., 2016). Whether p53 activates or restricts L1 expression could be affected by the used models and tools in the described studies.

Pluripotency transcription factors implicated in embryonic stem cell biology such as Oct4, Nanog, Sox2, KLF4, MYC, and LBP9 have functional binding sites in HERVH retrotransposon, and they support its RNA expression (Wang et al., 2014). Another paper

showed Oct4 and Sox2, but not Nanog can bind and activate HERVK retrotransposon expression in human embryonic carcinoma and HEK293T cells (Grow et al., 2015). Binding sites of CTCF, Oct4 and Nanog were enriched in retrotransposons and activated their transcription (Kunarso et al., 2010). CTCF binding sites were conserved across cell types and species, whereas Oct4 and Nanog binding sites had different enrichment profiles according to the cell type and origin (Kunarso et al., 2010). A recent comprehensive study used the Encyclopedia of DNA Elements (ENCODE) Chromatin Immunoprecipitation (ChIP)-sequencing datasets identified 138 transcription factors which bind L1 promoters (Sun et al., 2018). Among these factors, Myc and CTCF were shown to colocalize at the L1 promoter and regulate L1 differently (Sun et al., 2018). They are parts of a regulatory mechanism, where Myc represses L1, and CTCF promotes L1 activity (Sun et al., 2018).

Although the L1 5'UTR promoter region is prone to higher mutation rates than L1 ORF regions, evolutionary analysis showed conservation in transcription factors binding sites among human-specific L1 elements (Thornburg et al., 2006). Transcription factors in their regulation of retrotransposon expression are not isolated from other regulators that play roles in modulating retrotransposon activity in the cell. Each of these regulators is a part of different pathways that make up an interconnected network of factors controlling retrotransposon expression and activity.

1.4.2 Nuclear Regulation of Retrotransposons

The initiation of gene transcription is controlled by DNA methylation, a process carried out by DNA methyltransferase enzymes (DNMT) that act as regulatory switches for

epigenetic modifications required for normal mammalian development (Jones, 2012; Grandi et al., 2015). This modification involves a covalent addition of a methyl group to the fifth carbon of cytosine in cytosine-phosphate-guanine (CpG) sites that silences gene expression by altering the extent of DNA accessibility to transcription machinery (Smith and Meissner, 2013; Grandi et al., 2015). Most CpGs in the human genome are generally methylated, with less than 10% occurring in CG dense regions where they are referred to as CpG islands (Smith and Meissner, 2013). These islands are prevalent at the TSS of housekeeping and developmental regulator genes (Smith and Meissner, 2013). Retrotransposons contain about half of the CpG islands in the human genome (Fazzari and Greally, 2004; Grandi et al., 2015). When genes are expressed, CpG islands in their promoters undergo hypomethylation (Goodier, 2016). Consistent with that, when the L1 expression is typically suppressed, a CpG island adjacent to the L1 promoter in its 5'UTR is hypermethylated (Goodier, 2016). Evidence suggests that epigenetic silencing through DNMT mediated DNA methylation of CpGs has evolved to serve as a defence mechanism against retrotransposon activity (Goodier, 2016).

In addition to CpG DNA methylation, histone epigenetic modifications are used to regulate retrotransposon expression (Grundy et al., 2021). Histones are positively charged proteins used to package the human genome (Strahl and Allis, 2000). Their covalent modifications can change the accessibility of a particular stretch of DNA to transcription machinery (Strahl and Allis, 2000). However, these marks can induce or repress transcription, unlike DNA methylation (Strahl and Allis, 2000). The trimethylation of lysines 9 and 27 on histone 3 (H3K9me3 and H3K27me3) are two common suppressive

histone modifications associated with heterochromatin and frequently found on nucleosomes at TE loci (Dunham et al., 2012). These repressive histone modifications are deposited by multiple enzymes that form multimeric complexes (Grundy et al., 2021).

Retrotransposon transcription is silenced by a complex formed when Krüppel- associated box domain zinc finger proteins (KRAB-ZFPs) family proteins recruit KAP1/ TRIM28 (Ecco et al., 2017). KAP1 acts as a scaffold in this complex comprised of multiple DNA binding proteins, including SETDB1 (histone methyltransferase), NuRD (nucleosome remodeling and deacetylation) complex, HP1 (heterochromatin protein 1) and DNA methyltransferases (Ecco et al., 2017). The KRAB-ZFP/KAP1complex acts as a transcriptional repressor of retrotransposons by inducing heterochromatin formation in somatic cells and promoting DNA methylation in early embryonic cells (Ecco et al., 2017). Evidence suggests that by repressing retrotransposons, the KRAB-ZFP/KAP1 complex permits the normal development of embryonic cells and maintenance of transcriptional homeostasis. However, other evidence indicates the enrichment of activation marks on chromatin of KRAB-ZFP/KAP1 bound retrotransposons (Imbeault et al., 2017). Therefore, two models are suggested for retrotransposon transcription regulation by the KRAB-ZFP/KAP1 complex, an “arms race” model and a “domestication” model (Ecco et al., 2017). The first describes the KRAB-ZFP/KAP1 transcription repression function imposed on retrotransposons, and the second describes the co-evolution of both KRAB-ZFPs and retrotransposons to ensure that genetic homeostasis is in place (Ecco et al., 2017). The latter can exemplify the ability of our systems to adapt and deal with any defence breach.

1.4.3 Cellular Regulation of Retrotransposons

Some shared characteristics between retrotransposons and viruses can be considered outcomes of a similar origin as indicated by early phylogenetic studies (Xiong and Eickbush, 1990; Malik, 2000), which inferred that the RT sequence of an ancient non-LTR retrotransposon evolved to generate both gag-like and pol-like LTR retrotransposons and retroviruses subsequently (Xiong and Eickbush, 1990). Therefore, the cells have adapted and developed multiple defence lines against endogenous retrotransposons and exogenous viral infections (Goodier, 2016). As expected, some of the factors and mechanisms restraining retrotransposons in human cells are shared with the cellular antiviral response (Goodier, 2016).

1.4.3.1 Cytosine Deaminases

The family of cytosine deaminases (AID, APOBEC1, APOBEC2, APOBEC3 and APOBEC4) contains antiviral factors that restrict retrotransposons via multiple mechanisms (Esnault, 2006; Milewska et al., 2018). According to the family member, these enzymes act by deaminating cytosine to uracil within DNA and RNA molecules (Swanton et al., 2015). Of the most extensively studied, APOBEC3 includes seven sub-members in humans, A3A, A3B, A3C, A3D, A3F, A3G and A3H (Goodier, 2016). These restrict retrotransposons to varying extents, but A3A and A3B are considered the most effective (Goodier, 2016). The restriction mechanism of retrotransposons by members of APOBEC3 can be deamination dependent or independent (Feng et al., 2017). APOBEC3 deaminates the transiently exposed single-strand DNA generated during L1 integration in the first pathway, which triggers DNA degradation by the excision repair pathway and

restricts retrotransposition (Richardson et al., 2014). APOBEC3 can physically interact with L1 RT through the deamination-independent pathways and interfere with DNA polymerization during TPRT or target the RNP complexes for sequestration in cytoplasmic compartments such as stress granules (SGs) (Horn et al., 2014; Liang et al., 2016). The literature is unclear about the cellular conditions in which APOBEC3 utilizes either the deamination-dependent or -independent pathways (Feng et al., 2017).

1.4.3.2 Aicardi-Goutières syndrome associated genes

Other inhibitors of retrotransposons that are also involved in the anti-retroviral response include SAM domain and HD domain 1 (SAMHD1) and Three-prime repair exonuclease 1 (TREX1) (Goodier, 2016). These are two of seven genes linked to an early-onset autoinflammatory disorder mainly affecting the brain called Aicardi-Goutières syndrome (AGS) (Benitez-Guijarro et al., 2018). SAMHD1 acts as a dGTP-activated deoxynucleoside triphosphate triphosphohydrolase (dNTPase), and TREX1 is a 3'-to-5' DNA exonuclease (Goodier, 2016).

SAMHD1 degrades deoxynucleoside triphosphates (dNTPs) to regulate intracellular nucleotide levels, and it can block retroviruses' and retrotransposons' replication in arrested cells and cell cultures, respectively (Herrmann et al., 2018). By providing a crystal structure of homotetrameric SAMHD1, its mechanistic roles in restricting retroelements started to be clarified (Zhu et al., 2013). SAMHD1 regulates L1 in a similar but alternative mechanism to its HIV-1 regulation. Forming a dGTP-triggered tetramer was demonstrated to be essential for SAMHD1-mediated HIV-1 restriction and L1

limitation (Zhu et al., 2013). However, SAMHD1 regulation of L1 was shown to be independent of dNTP hydrolase activity (Herrmann et al., 2018). To restrict L1 effectively, SAMHD1 is phosphorylated at threonine 592 (T592), and it interacts directly with ORF2p in L1 RNP complexes (Herrmann et al., 2018). This restriction relies on an enzymatically functional HD domain and the SAMHD1 allosteric GTP-binding sites (Herrmann et al., 2018).

The primary function of TREX1 is to deplete damaged DNA by acting on both ssDNA and double-stranded DNA (dsDNA) (Li et al., 2017). Although it was demonstrated to target the reverse transcribed DNA of retrotransposons to limit their capacity to insert into the genome and elicit type I interferon (IFN) responses (Stetson et al., 2008), TREX1 inhibits L1 activity by an exonuclease-independent mechanism (Li et al., 2017). It interacts with ORF1p to change its subcellular localization, triggers its depletion and causes a reduction in L1 retrotransposition subsequently (Li et al., 2017).

1.4.3.3 Small RNA Silencing Pathways

Another mechanism regulating the expression of retrotransposons is double-stranded RNA (dsRNA) that induces sequence-specific gene silencing pathways (Malone and Hannon, 2009). When processed into small RNAs, 20- 30 nt in length, with the help of RNaseIII enzyme Dicer, dsRNA initiates a gene silencing process that can occur in different complexes (Carthew and Sontheimer, 2009). RNA interference (RNAi) is one of the cellular strategies employed to restrict retrotransposon's activity (Malone and Hannon, 2009). RNAi can act at the post-transcriptional level by inducing target RNA cleavage or

degradation via a ribonuclease-containing RNA-induced silencing complex (RISC) or at the transcriptional level by causing epigenetic modifications (Malone and Hannon, 2009; Goodier, 2016). Piwi-interacting RNA (piRNA) is another class of small RNAs processed independently of DICER, which is slightly longer than siRNAs (Siomi et al., 2011). They associate with PIWI proteins to form a piRNA-induced silencing complex (piRISC), which protects genome integrity from retrotransposon threats by silencing them in the germline and gonadal somatic cells (Siomi et al., 2011). piRISC allows PIWI proteins to specifically recognize and cleave retrotransposon transcripts by PIWI guided by piRNAs whose sequences are antisense to retrotransposon RNAs (Siomi et al., 2011). PIWI proteins and piRNAs can also mediate CpG DNA methylation of retrotransposon promoters to restrict their activity at the transcriptional level (Siomi et al., 2011).

1.4.3.4 Other Regulators

Evidence suggests multiple other cellular molecules restrict retrotransposon activity; however, their mechanism of action is not well defined (Goodier, 2016). In this section, three more regulators involved in cellular antiviral response are covered, MOV10, RNase L and ZAP.

1.4.3.4.1 MOV10

Moloney leukemia virus 10 (MOV10) is a member of the UPF1-like superfamily1 of ATP-dependent RNA helicases (Choi et al., 2018). MOV10 was first identified as a restriction factor against Moloney leukemia virus infections (Choi et al., 2018). MOV10 is implicated in the inhibition of retrotransposon replication in human cellular models (Choi et al., 2018). MOV10 is suggested to cause sequestration of L1 RNP and

degradation of L1 RNAs since it colocalizes with ORF1p and components of the RISC pathway in SGs and cytoplasmic processing bodies (P-bodies) (Goodier, 2016; Choi et al., 2018). Recent evidence has demonstrated a direct interaction between MOV10 and Ribonuclease H2 (RNase H2) in an RNA-dependent manner (Choi et al., 2018). This interaction inhibits L1 retrotransposition by degrading L1-derived RNA-DNA hybrids (Choi et al., 2018).

1.4.3.4.2 RNase L

The IFN inducible endoribonuclease, RNase L, inhibits viral replication by binding single-stranded regions of viral and cellular RNAs and promoting their cleavage (Goodier, 2016). Viral dsRNA stimulates oligoadenylate synthetase that uses ATP to synthesize 2',5'-oligoadenylate (2-5A) (Goodier, 2016). These 2-5A molecules bind latent RNase L and activate it to dimerize (Goodier, 2016). Upon its prolonged activation, RNase L promotes the death of virus-infected cells by inducing autophagy and apoptosis pathways (Goodier, 2016). In cell culture assays, RNase L was demonstrated to restrict mouse IAP and human L1 retrotransposition suggesting a role in targeting their RNA for degradation; however, the exact mechanism of inhibition is still open for further clarification (Zhang et al., 2014).

1.4.3.4.3 ZAP

The zinc-finger antiviral protein (ZAP) is a member of the PARP family of proteins and another restrictor of retrotransposons activity (Moldovan and Moran, 2015). Due to alternative splicing, it exists in two spliced isoforms, long and short. The long isoform of ZAP (902 amino acids long) contains a defective C-terminal PARP-like domain (Kerns et al., 2008). The short isoform of ZAP (699 amino acids long) lacks the C-terminal PARP-

like domain and is induced by IFN (Kerns et al., 2008). ZAP N-terminal CCCH-type zinc finger domain (exists in both isoforms) binds and targets several RNA viruses' transcripts for degradation in cytoplasmic granules (Zhu et al., 2011; Moldovan and Moran, 2015). The two ZAP isoforms exhibited similar restriction activity against retrotransposons in cell culture assays (Moldovan and Moran, 2015). The evidence of ZAP preventing the accumulation of L1 mRNA in the cytoplasm and being colocalized with ORF1p and L1 RNA in SGs indicates that ZAP is likely recruiting RNA degradation proteins to retrotransposon transcripts (Moldovan and Moran, 2015).

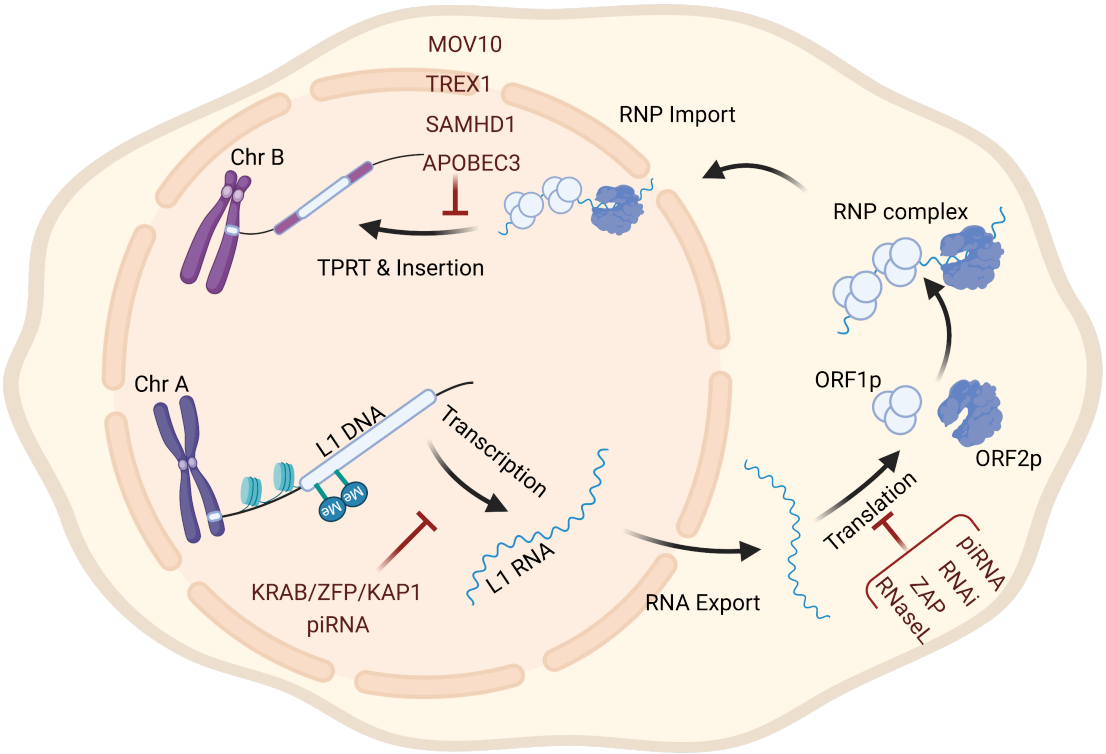


Figure 1-3 Retrotransposon levels of regulation throughout its life cycle.
Regulation of retrotransposons activity can occur at the transcriptional level by histone modification or DNA methylation. At the post-transcriptional level by targeting RNA for degradation. At the genomic insertion level, by interfering with RNP complexes integrity or inhibiting TPRT. Illustrations were created with BioRender.

1.4.4 Upon Defence Failure

Retrotransposons were long considered genetic parasites that evolved to stay active in the germline to be transmitted along to the next generation but get inactivated in the soma as a means of protection against their harmful effects (Goodier, 2016). With all the above-described mechanisms to restrict retrotransposons, active insertions still occur and are detected in somatic cells and tissues (Goodier, 2016). L1 retrotranspositions were identified in neuronal precursor cells (Muotri et al., 2005), can occur during early human embryonic development (van den Hurk et al., 2007) and in various types of cell lines when tagged L1 construct is employed (established in the field). However, limited data is available on whether retrotranspositions occur in normal somatic adult tissues other than the brain (Coufal et al., 2009; Evrony et al., 2012; Upton et al., 2015) except for a few findings of identified somatic insertions in hepatocytes and DNA of esophagus, stomach and colon, which may have occurred during embryogenesis (Doucet-O'Hare et al., 2015, 2016; Ewing et al., 2015). This lack of evidence could be related to the non-dominating occurrence of somatic insertions in few cells, which is challenging to identify in whole tissue sequencing. On the other hand, multiple numbers of new somatic insertions have been identified in different types of tumor tissues of epithelial origin at varying frequencies (Lee et al., 2012; Scott and Devine, 2017). These insertions are characterized by having more 5' truncations and existing with less dependence on L1 encoded EN cleavage than germline insertions (Solyom et al., 2012).

Retrotransposons deregulation (due to the failure of restriction mechanism) has been associated with the onset of aging and several disease conditions (Ishak et al., 2018). One

of the first links of retrotransposons to disease was made back in the 1980s when Haig Kazazian identified disruptive L1 insertions in the factor VIII gene of a subset of hemophilia A patients (Kazazian et al., 1988). About the same time, a gene-activating rearrangement caused by L1 insertion at the *MYC* oncogene was reported in a breast cancer patient (Morse et al., 1988). After that, a large body of evidence revealed and is still revealing retrotransposons as essential contributors to the course of tumorigenesis, especially with the global epigenetic dysregulation that characterizes tumorigenesis (Sharma et al., 2010).

1.4.5 Association of retrotransposons to cancer

The association of retrotransposon activity to tumorigenesis was identified in recent times following some exciting observations in tumors. One of these was when Miki *et al.* detected a *de novo* L1 insertion in an exon of one allele of the tumor suppressor APC in colon cancer but not in normal colon tissues from affected individuals (Miki et al., 1992). More than two decades later, another novel L1 insertion was found to disrupt the other allele of the *APC* gene, which contributed to colon cancer initiation by the classical "two-hit" model (Scott et al., 2016). *ST18* (Suppression of Tumorigenicity18) and PTEN genes were two other tumor suppressors interrupted by new L1 insertions in hepatocellular carcinoma and endometrial cancer, respectively (Shukla et al., 2013a; Helman et al., 2014). Retrotransposon activity in cancer is a subject that has attracted many research groups trying to understand the nature of the association. By immunolabeling a panel of 1027 different human cancers in a collection of tissue microarrays for L1 (ORF1p), L1

protein expression was demonstrated to be a common characteristic among high-grade malignant cancers of epithelial origin (Rodić et al., 2014).

On the other hand, ORF1p was infrequently detected in the early stages of tumorigenesis and was absent from normal somatic tissues (Rodić et al., 2014). With the advances made in high-throughput sequencing and bioinformatic analysis, researchers have been using sophisticated tools and technologies to study retrotransposon activity in the context of cancer from a deeper perspective. By introducing the high-throughput L1-Sequencing assay, Iskow *et al.* could identify a hypomethylation signature that characterized lung tumors, which made them more L1-permissive and have a higher frequency of L1 somatic insertion than the brain tumors included in the study (Iskow et al., 2010). Tumors of epithelial origin such as colorectal, prostate, and ovarian cancers showed more pronounced L1 activity than brain and blood cancer types by performing single-nucleotide resolution analysis of TE insertions in whole-genome sequencing datasets (Lee et al., 2012). In agreement with these findings, colorectal and lung cancers were the most frequently affected by L1 somatic insertions that exhibited hypomethylated promotors by tracking down the L1 insertion sources via identifying 3' transductions (Tubio et al., 2014). However, these insertions demonstrated minimal to no effect on the course of tumorigenesis (Tubio et al., 2014). This evidence could be related to the variations caused by different cancer types, stages, or received therapy. Despite this, the association of retrotransposon activity to carcinogenesis is well established in the literature. In addition, research is still ongoing to answer exciting questions, including tumor type specificity, insertion timing throughout tumor evolution and the impact of retrotransposon activity on

cancer biology and treatment. The recent advances in these areas are to be discussed further in the following sections.

1.4.5.1 Retrotransposons and immune response

As mentioned earlier, some of the retrotransposon regulation mechanisms are used similarly to protect cells from exogenous viral infections. When nucleic acids of foreign origin are detected in the cytoplasm by endosomal or pattern recognition receptors (PRR), an IFN-driven immune response is initiated to eliminate the affected cell populations (Ishak et al., 2018). The cell is equipped with a heterogeneous group of PRRs that includes, but not limited to: Toll-Like Receptors (TLR3, TLR7, TLR8, and TLR9), the RNA sensors RIG-I (Retinoic-acid-Inducible Gene I) and MDA5 (Melanoma Differentiation-Associated protein 5), and the DNA sensors cGAS (cyclic GMP-AMP Synthase) and AIM2 (Absent in Melanoma 2) (Roers et al., 2016).

Specific criteria determine which nucleic acid is sensed by each PRR, including location, nucleic acid sequence pattern, and a threshold quantity (Roers et al., 2016). The nucleic acid binding domains of TLRs face the lumen of endosomal compartments, and the other PRRs are present in the cytoplasm (Roers et al., 2016). TLR3 binds dsRNA of >40 bp size, TLR7, 8 bind fragmented RNA with unmodified nucleosides, and TLR9 binds ssDNA of > 11 nt size with a high affinity to unmethylated cytosine CpG motif (Roers et al., 2016). RIG-I binds >20 bp dsRNA with blunt end conformation, MDA5 binds >1-2 Kb dsRNA, cGAS binds dsDNA of >20-40 bp size, and AIM2 binds dsDNA of >50-80 bp size (Roers et al., 2016). The quantity of detected nucleic acid can be affected by two

factors, the increased supply that causes the accumulation of nucleic acids and the defective mechanisms of their clearance.

The failure of one or more of the retrotransposon restriction mechanisms (described in the previous section) due to aging, tumorigenesis, or autoimmune disease can result in retrotransposon activation. This activity promotes dsRNA or dsDNA (sequences of different sizes and motifs) release into the cytoplasm and their detection by PRRs (Yu et al., 2012; Chiappinelli et al., 2015; Roulois et al., 2015; Shen et al., 2015). During aging onset, retrotransposons could evade their restriction mechanisms and get activated (De Cecco et al., 2019). This activation of L1 during cellular senescence triggered the release of L1 dsDNA in the cytoplasm and promoted type I IFN response and sterile inflammation (De Cecco et al., 2019). Multiple examples are available for retrotransposon activation in cancer. By analysing TCGA RNA sequencing data, specific HERV elements were highly enriched in tumor samples compared to their normal counterparts, and this enrichment was associated with increased immune response (Rooney et al., 2015). Another evidence showed that cytosolic ssDNA and dsDNA in several tumor cell lines were mainly retrotransposon-derived and associated with the cGAS activated type I IFN response (Shen et al., 2015). Activating HERV expression using DNMT inhibitors (DNMTi) in cancer cells triggered cytosolic dsRNA release and MDA5 stimulated immune response (Chiappinelli et al., 2015). In addition, expressing HERV sequences in TLR3, TLR7, and TLR9 triple-deficient mice failed to induce sufficient immune response that resulted in their development of T cell acute lymphoblastic leukemia and their early death (Yu et al., 2012). Blood samples from individuals with the autoimmune disease SLE, Systemic

Lupus Erythematosus, were enriched in *Alu* RNA associated with high levels of type I IFN response (Hung et al., 2015). Although the triggers of retrotransposon activation in the disorders mentioned above may differ, they mostly shared the activation of the IFN response as a consequence.

Modulating retrotransposon expression to regulate its associated immune response is an attractive therapeutic strategy under current investigation (Ishak et al., 2018). Retrotransposon activation may be achieved using drugs that inhibit DNA methylation (DNMTi) or histone deacetylation (HDACi) (Ishak et al., 2018). The use of DNA-hypomethylating agent 5-azacitidine (AZA) in colon and ovarian cancer cell models was associated with increased expression of HERV and L1 RNA (Chiappinelli et al., 2015; Desai et al., 2017). HERV expression was linked to regulatory T cells tumor infiltrates and predicted cytolytic activity in AZA-treated cells (Desai et al., 2017). In contrast, L1 expression correlated with p53 status and predicted AZA drug sensitivity (Desai et al., 2017). A dinitroazetidine derivative (RRx-001) is another hypomethylating drug, less toxic than AZA and is currently in phase II clinical trials (Zhao et al., 2017). RRx-001 was shown to induce antitumorigenic effects by activating the expression of HERV and IFN-responsive genes consequently (Zhao et al., 2017). Similarly, treating colon cancer cells and tumor organoids with another derivative of hypomethylating agent (5-aza-2′-deoxycytidine) was sufficient to induce a growth-inhibiting immune response by triggering retrotransposon expression (Roulois et al., 2015; Saito et al., 2016). Interestingly, the combination of DNMTi and HDACi selectively induced LTR retrotransposons with higher efficiency than using each drug individually (Brocks et al.,

2017). By analyzing the treatment-activated TSS of these elements, they were *de novo* induced from non-annotated TSS (Brocks et al., 2017). This activation resulted in chimeric products with predicted abnormal or immunogenic functions (Brocks et al., 2017).

In addition, some targeted cancer therapeutics and chemotherapeutic agents were shown to activate retrotransposon expression in cancer cells (Goel et al., 2017; Guler et al., 2017). The use of cyclin-dependent kinases 4 and 6 (CDK4/6) inhibitors in breast cancer repressed DNMT1 and caused activation of repeat elements, including retrotransposons (Goel et al., 2017). This activation promoted cytotoxic T-cell-mediated clearance of tumour cells and increased tumour immunogenicity (Goel et al., 2017). However, some cells within a heterogeneous population of cancer may develop adaptation mechanisms that allow their survival in the challenging conditions of the tumor microenvironment (Guler et al., 2017). These cells could modulate retrotransposon expression with lethal drug exposures by maintaining epigenetic repression on them (Guler et al., 2017). This evidence suggests that combining HDACi with other targeted therapeutics may enhance their efficacy in treating cancer (Ishak et al., 2018).

The examples mentioned above support the notion that retrotransposon activation in tumors may contribute to their turning into ‘hot tumors,’ which are inflamed and T-cell infiltrated tumors (Galon and Bruni, 2019). In such a microenvironment, the antitumor immune response will reduce the tumor burden and sensitize it to other types of targeted therapies and immunotherapy (Galon and Bruni, 2019)

Retrotransposon activity in cancer is probably occurring in specific tumor types over others (Lee et al., 2012; Scott and Devine, 2017), and it is not clear yet whether this is related to a more vigorous immune defence or a higher level of cellular adaptation (Burns, 2017). The following section will discuss retrotransposon activity in cancer for a better understanding and evaluation of recent evidence.

1.5 Retrotransposons and Cancer

Retrotransposons might contribute to disease through their high potential to evade defence mechanisms, propagate and insert in other parts of the genome causing structural and functional impacts (discussed in the previous sections). Retrotransposons are strongly induced during spermatogenesis and oocyte development, and their overactivation contributes to the selective elimination of 80% of maturing oocytes (Tharp et al., 2020). Similarly, the significant evolutionary drift caused by transposable elements may also occur during tumor evolution and might contribute to cancer outcomes (Goodier, 2016).

1.5.1 Tumors prone to retrotransposons activity

The underlying mechanism causing some tumors to have more retrotransposon activity than others is not well understood (Burns, 2017). In general, the vast majority of available evidence suggests that retrotransposon insertions are occurring more frequently in tumors of epithelial origin (Lee et al., 2012; Scott and Devine, 2017). L1 retrotranspositions were detected at the highest levels in lung and colorectal tumors (Tubio et al., 2014; Ewing et al., 2015; Scott et al., 2016). Esophageal, pancreatic, head and neck, uterine, ovarian, gastric, and prostate cancers displayed moderate to high levels of L1 activity (Helman et al., 2014; Tubio et al., 2014; Doucet-O'Hare et al., 2015; Ewing et al., 2015; Rodić et al.,

2015; Tang et al., 2017; Rodriguez-Martin et al., 2020). In contrast, lower levels of L1 activity were detected in breast, bone, liver, kidney, and testicular cancers (Shukla et al., 2013a; Helman et al., 2014; Tubio et al., 2014; Ewing et al., 2015). Also, tumors with high L1 activity showed a corresponding increase in L1 protein expression (Rodić et al., 2014), except for ovarian cancer samples that demonstrated moderate L1 activity and high expression of ORF1p (Rodić et al., 2014).

This preference of retrotransposons for distinct tumor types could be related to a range of transcription factors activated in specific cell types over others. Activation of transcription factors in epithelial tumors might modulate retrotransposon expression and activity. For example, epithelial tumors such as breast, colorectal, prostate and cervical cancers were characterized by Oct1 (Octamer transcription factor 1, POU2F1) protein up-regulation (Maddox et al., 2012; Obinata et al., 2012; Xiao et al., 2014; Wang et al., 2016). Oct1 was demonstrated to control stem cell phenotypes in normal and tumor cells (Maddox et al., 2012). In epithelial cells, high Oct1 protein expression was spatially correlated with stem cell niches and increased expression of stem cell markers such as ALDH1 (Maddox et al., 2012). Transcription factors like Oct1 may be playing a role in the ability of epithelial cells to be reprogrammed and become pluripotent stem cells (Carreira et al., 2014). These reprogrammed differentiated cells may be more disposed to L1 retrotransposition than other populations of cancer cells (Carreira et al., 2014). In addition, Oct1 binding sites were associated with L1 and *AluY* DNA methylation in a study that involved hundreds of placental tissues and investigated mechanisms affecting fetal development (Wilhelm-Benartzi et al., 2012).

In general, there is no definitive interpretation from available evidence for the prevalence of L1 activity in some tumor types. However, the recent genome-wide studies can provide the field with more in-depth information about retrotransposon activity and tumor genomes. One of the thoughts to be considered when looking to differentiate between tumors is the timing of retrotransposon insertions during tumor evolution, which could be characteristic to certain tumors over others and can provide some elucidations to some of the open questions (to discuss in the next section).

1.5.2 Driver versus passenger mutation

Due to cancer heterogeneity and the transportable nature of retrotransposons, it is challenging to differentiate tumor-driver retrotransposon insertions from tumor-passenger mutations. However, the advances in molecular biology and bioinformatic analysis tools made it possible to characterize *de novo* insertions throughout tumor development stages.

One of the fundamental advancements in this regard was identifying active full-length L1 sequences that are retrotransposition-competent in the human genome, and each individual is estimated to have 80-100 elements of them (Brouha et al., 2003). Based on collected evidence from cancer patients, some retrotransposon insertion events were detected, interrupting tumor suppressors, which supports considering them as tumor driver mutations (Burns, 2017). While other insertions were detected at later tumor stages or in metastatic tumors, supporting the notion of being passenger mutations (Burns, 2017).

For example, the early report of gene-activating rearrangement within *MYC* oncogene in breast cancer and the reports of disruptive insertions detected in APC tumor suppressor in colon cancer patients suggest these insertions as cancer 'driver' mutations (Morse et al., 1988; Miki et al., 1992; Scott et al., 2016). Also, a recent comprehensive study analyzed 2,954 cancer genomes from 38 different tumor subtypes, identified several oncogenic roles of L1 retrotransposition that include structural variations encompassing tumor suppressor genes and triggered amplification of oncogenes (Rodriguez-Martin et al., 2020). On the other hand, somatic L1 insertions acquired during tumor evolution in different cancer types were widely distributed in intronic and intergenic regions without affecting tumor suppressor or oncogenic loci (Tubio et al., 2014; Rodić et al., 2015). These insertions are likely 'passenger' rather than 'driver' mutations (Tubio et al., 2014; Rodić et al., 2015).

Although retrotransposon insertion sites share common characteristics such as preferred DNA target motif and nucleosome content, they are dispersed across all chromosomes with no specific preference (Sultana et al., 2019). It is still early to have a definitive (cause and effect) answer about what predicts the sites of genomic insertions and at what time of tumor evolution they occur. In addition, cancer is a complex system that involves multiple genetic and environmental factors that can play a role in activating retrotransposons (Sinibaldi-Vallebona et al., 2011). Therefore, studying retrotransposon activity in the context of cancer requires considering the underlying mechanisms and contributing factors that might play a role in activating them.

1.5.3 Impact of retrotransposons on tumorigenesis

L1 activation in cancer is strongly associated with general hypomethylation status, alleviating tight control over this transposable element (Hur et al., 2014; Tubio et al., 2014). However, the principal causes of this hypomethylation are primarily unknowns (Xiao-Jie et al., 2016). Some factors and pathways were suggested to play essential roles in this hypomethylation without knowing the exact mechanisms, such as oxidative stress, chronic inflammation, p53 deficiency, and Wnt/β-catenin signaling pathway (Xiao-Jie et al., 2016). The consequences of the described L1 activation in cancer are limited to a few 'driver' mutations that can be considered tumor-initiating; however, the majority of L1 insertions are acquired throughout tumor evolution in non-coding DNA regions and posing occasional effects on the course of tumorigenesis (Lee et al., 2012; Tubio et al., 2014; Rodriguez-Martin et al., 2020). The research is still ongoing to understand the impact of the L1 activity on tumor development, exploring L1-related functions such as RT activity.

Throughout their evolutionary timeline, significant retrotransposon-related activities at the genomic and cellular levels are attributed to their RT (Mathias et al., 1991). However, retrotransposon genomic insertions in cancer have drawn considerable attention beyond the attention given to retrotransposon RT activity (Sciamanna et al., 2016). RT activity was shown to increase in cancer, and by using anti-retroviral non-nucleoside reverse transcriptase inhibitors (NNRTIs) such as efavirenz and nevirapine, RT activity could be reduced significantly by inducing conformational changes in the enzyme that prevented its activity (De Clercq, 1993; Contreras-Galindo et al., 2008). NNRTIs were demonstrated

to effectively reduce tumor growth by decreasing cellular proliferation and promoting differentiation (Mangiacasale et al., 2003; Landriscina et al., 2007). The effect of inhibiting RT using NNRTIs was similar to the L1 siRNA suppressing effect; therefore, they were assumed to target L1 activity (Sciamanna et al., 2005). Other lines of evidence suggest that another class of RT inhibitors, nucleoside reverse transcriptase inhibitors (NRTIs), are capable of inhibiting L1 activity and having anticancer effects in cells (Jones et al., 2008; Carlini et al., 2010). This evidence suggests that L1 encoded RT serves as a potential marker for diagnostic purposes and an exciting target for therapeutic intervention. However, further work is still required to understand the exact mechanism of the observed effect of RT inhibitors on cancer (Rodić and Burns, 2013). This effect could be related to a pathway other than the control of L1 activity such as telomerase RT, since that NRTIs, but not NNRTIs, were shown to inhibit telomerase RT *in vitro* (Hukezalie et al., 2012).

Tumor-derived extracellular vesicles (EVs) were shown to be enriched in retrotransposons RNA and involved in horizontal transfer of retrotransposons to normal cells along with other oncogenic sequences (Balaj et al., 2011; Kawamura et al., 2019). This evidence suggests that EVs facilitate the release and transfer of retrotransposons to other cells, contributing to tumor evolution or metastasis (if derived from tumor cells). Also, retrotransposon RNA transfer can influence the transcriptional and post-transcriptional regulation of recipient cells. For example, APOBEC3 members were activated in response to the increased L1-derived RNA transcripts in recipient cells after the EVs transfer (Kawamura et al., 2019). Research is still ongoing to characterize the content of EVs on

the way to use them as non-invasive tools for cancer detection. The increased expression of retrotransposons in EVs derived from tumor cells compared to those derived from normal cells (Balaj et al., 2011) can serve as a valuable biomarker for diagnostic purposes. Current evidence on retrotransposons containing EVs highlights the impact of retrotransposon RNA, suggesting a way retrotransposon can contribute to tumorigenesis and opens opportunities to study retrotransposon RNA expression activation in cancer. Studies to characterize the retrotransposon containing EVs origin, biogenesis and destination in cancer patients are currently needed to understand their potential fully.

1.6 Retrotransposon Activity in Ovarian Cancer

A high death-to-incidence rate characterizes ovarian cancer compared to other gynecological cancers (Lheureux et al., 2019). The latest cancer statistics worldwide indicate that about 66% of newly diagnosed ovarian cancer patients die of the disease annually (Ferlay et al., 2021). That is a significant percentage compared to the 30% death-to-incidence rate in breast cancer (Ferlay et al., 2021). The advanced stage of ovarian cancer at the time of diagnosis is considered one of the main contributors to this high death-to-incidence rate (Lheureux et al., 2019). Only 29% of women diagnosed with late-stage ovarian cancer have a 5-year survival rate (Doherty et al., 2017). In comparison, about 92% of women with early-stage presentation have a 5-year survival rate (Reid et al., 2017). Ovarian cancer's different types can contribute to disease presentation and survival rates consequently (Ferlay et al., 2021). In addition to the histological appearance, ovarian cancer subtypes differ in their cells of origin, molecular alterations, clinical characteristics

and epidemiologic considerations (Ferlay et al., 2021). Differences between subtypes can have clinical implications in ovarian cancer diagnosis and treatment.

In the clinic, two criteria are taken into consideration at the time of ovarian cancer diagnosis: tumor type (according to the World Health Organization classification) and tumor stage (according to the International Federation of Gynaecology and Obstetrics (FIGO) classification) (Meinhold-Heerlein et al., 2015). Ovarian tumors can be serous, endometrioid, clear cell, seromucinous, mucinous, or Brenner subtype based on histopathological and molecular characteristics (Kurman et al., 2014).

1.6.1 Ovarian cancer subtypes

Serous tumors include both low-grade serous and high-grade serous subtypes. They were long thought to originate from the ovarian surface epithelium until recently (Lheureux et al., 2019), when accumulating evidence suggested they can also evolve from an extra-ovarian origin, specifically from the fallopian tube (fimbriae) epithelium closely associated with the ovarian surface cells (Brewer et al., 2005; Crum et al., 2007; Lee et al., 2007; Auersperg, 2013). High-grade serous tumors are the most prevalent and the most aggressive among other ovarian cancer subtypes (Lheureux et al., 2019). *TP53* mutations, deletions or inactivation characterize at least 96% of high-grade serous tumors (Bell et al., 2011; Chien et al., 2015). These tumors also bear lower prevalence mutations in other genes such as *BRCA1, BRCA2, NF1, RB1,* and *CDK12* (Bell et al., 2011). About half of high-grade serous tumors are defective in HR repair (Bell et al., 2011). Besides, these tumors generally have altered NOTCH, FOXM1, RB, and PI3K/RAS signaling pathways

(Bell et al., 2011). On the other hand, low-grade serous ovarian cancer is a rare subtype that accounts for 4.7% of serous types and 2% of all ovarian epithelial tumors (Matsuo et al., 2018). P53 immunostaining is required to differentiate between low-grade serous and high-grade serous tumors (Meinhold-Heerlein et al., 2016). MAPK pathway signaling plays a crucial role in low-grade serous tumorigenesis (Slomovitz et al., 2020). Mutations in genes such as *KRAS, BRAF, ERBB2,* and *NRAS* are predominant in low-grade serous tested samples (Jones et al., 2012; Hunter et al., 2015; Etemadmoghadam et al., 2017).

Endometrioid and clear cell ovarian tumors are uncommon ovarian tumors that are predominantly invasive (Meinhold-Heerlein et al., 2016). These tumors likely evolve from foci of endometriosis (Bell, 2005; Wiegand et al., 2010), abnormal growth of endometrial tissue outside the uterine cavity (Parasar et al., 2017). Endometrioid tumors have multiple subtypes (including tumors with squamous or mucinous differentiation, secretory transformations, oxyphile type, or germ cell-like patterns) (Lu and Chen, 2014). Endometrioid tumors frequently harbor mutations in genes such as *KRAS, PIK3CA, PTEN, CTNNB1, ARID1A,* and less prevalence in *TP53* (Cybulska et al., 2019). Clear cell ovarian tumors share characteristics with endometrioid tumors such as recurrent *ARID1A* and *PIK3CA* mutations and infrequent *TP53* mutations (Shih-Chu Ho et al., 2001; Kuo et al., 2009; Wiegand et al., 2010).

The seromucinous tumors are a more recently described subtype of ovarian cancer with little details about their origins, but because of their ARID1A mutations, they are closer to endometrioid than to serous subtype (Meinhold-Heerlein et al., 2015, 2016). The 2014

World Health Organization classification considered the invasive mucinous subtype of ovarian tumors as metastasis of extragenital (gastrointestinal) malignancy (Meinhold-Heerlein et al., 2015, 2016). Brenner tumors are rare and can be benign, borderline, or malignant tumors of transitional or urothelial epithelium origin (Ricotta et al., 2021). Figure 1-4 below summarizes the main ovarian cancer subtypes (likely originate from the female reproductive system) and their common associated mutations.

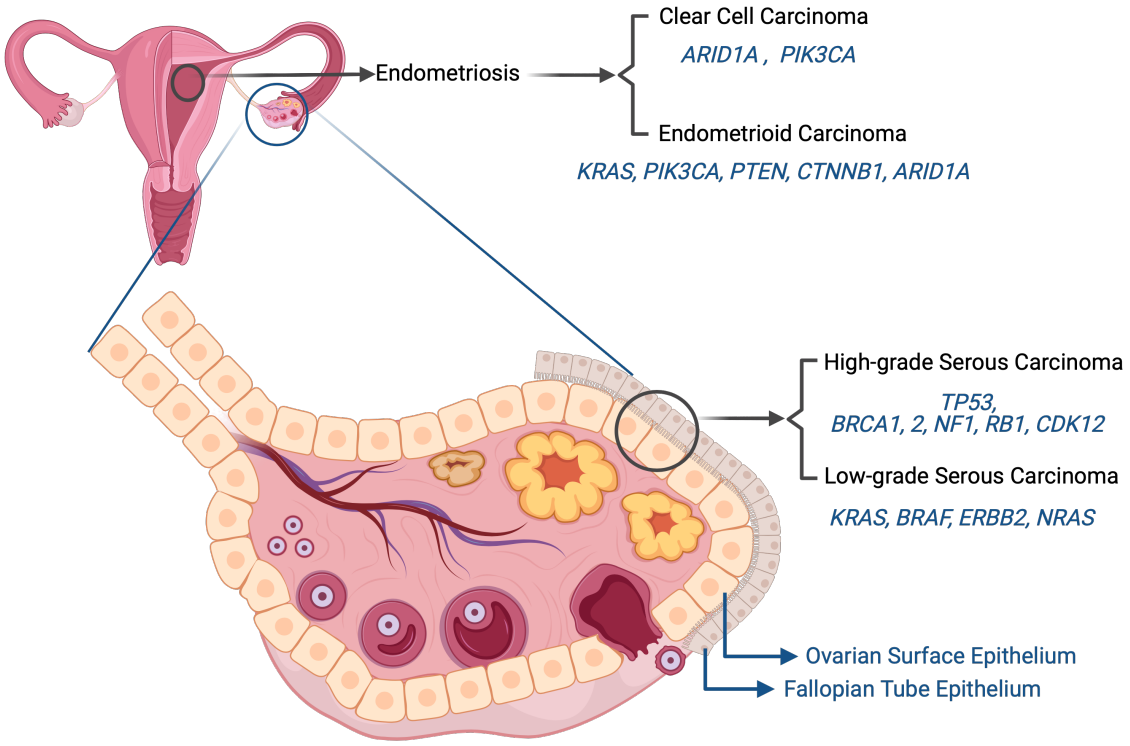


Figure 1-4 Ovarian cancer subtypes, their origin, and their common genes mutations.
Illustrations were created with BioRender.

1.6.2 Ovarian cancer stages

Based on their location at diagnosis, ovarian tumors are categorized into specific groups (stage I-IV) (Meinhold-Heerlein et al., 2015). In stage I, tumors are limited to the ovary or Fallopian tube. This stage underlies three subdivisions (IA, IB, and IC) in which IA describes a tumor when confined to one ovary or one tube with an intact capsule surrounding it. Stage IB describes the tumor when it involves both ovaries/tubes within intact capsules. Stage IC describes the tumor if it is limited to one or both ovaries or tube(s) with one of the following conditions: ruptured tumor capsule either intraoperatively, preoperatively, tumor exists on the ovarian, tubal surface, or if malignant cells found in the ascites or peritoneal cytology (Meinhold-Heerlein et al., 2015). Stage II includes tumors spread within the pelvis and is subdivided into stages IIA and IIB. IIA describes a tumor when it extends to the uterus, fallopian tubes, and ovaries, and IIB describes a tumor when it spreads to other pelvic intraperitoneal tissue. Stage III represents an advanced stage where tumor spread extends to outside the pelvis. This stage underlies four sub-stages (IIIA1, IIIA2, IIIB, IIIC). Stages IIIA1 and IIIA2 describe tumors when metastasis involves retroperitoneal lymph nodes only or microscopic extra-pelvic peritoneal metastases, respectively (Meinhold-Heerlein et al., 2015). Stage IV defines all tumors with distant metastases and encloses two subdivisions IVA and IVB. Stage IVA, when pleural effusion shows cancer positive cytology. Stage IVB describes liver metastases and metastases to extra-abdominal organs (Meinhold-Heerlein et al., 2015).

1.6.3 Treatment options for ovarian cancer patients

Despite significant advances in classifying ovarian tumors, the clinical care model uses a standard treatment approach for all classes, which might contribute to the unimproved high death-to-incidence rates of ovarian cancer remaining consistent over the years (Doherty et al., 2017). Ovarian cancer standard management includes a primary debulking surgery (aims to reduce as much as possible of the tumor's volume) followed by chemotherapy consisting of Carboplatin (platinum-based) and Paclitaxel (Bell et al., 2011). If the chemotherapy is followed by a debulking surgery (referred to as interval surgery), the approach is called neoadjuvant therapy which presents an alternative management approach (Kehoe et al., 2015). Patients with multiple comorbidities, large tumors or advanced disease stage (IV) can benefit from following this alternative approach by reducing surgical morbidity (Worley et al., 2014; Rauh-Hain et al., 2017).

In the last ten years, two clinical trials (EORTC55971 and CHORUS) indicated no difference between the primary surgery and the interval surgery (secondary surgery performed after neoadjuvant chemotherapy) based on the enrolled patients' overall survival except for stage IV patients (Vergote et al., 2010; Kehoe et al., 2015). A recent clinical trial (JCOG0602) indicated the absence of advantage of using neoadjuvant in stage IV patients and did not confirm the findings of the EORTC55971 and CHORUS trials (Onda et al., 2020). TRUST trial is ongoing to evaluate the incomparable findings of the three trials mentioned above (Reuss et al., 2019). Nevertheless, maximal (removing the maximum amount of tumor) debulking surgery remains the standard of ovarian cancer patient care regardless of the chemotherapy timing (Kurnit et al., 2021).

There is an inverse correlation between the patients' survival and the amount of residual tumor tissue (Ataseven et al., 2016). Therefore, the purpose of debulking surgery is the complete resection leaving no macroscopic remaining tumor tissue or at least leaving less than 1cm residual disease if complete cytoreduction is not feasible (Raspagliesi et al., 2013; Armstrong et al., 2021). Primary surgery of most patients (with suspected ovarian cancer) generally includes bilateral salpingo-oophorectomy (removal of both ovaries and fallopian tubes), hysterectomy (removal of the uterus if present), infracolic and infragastric omentectomy (removal of omentum, a flat adipose tissue covering intra-peritoneal organs) (Di Nicola, 2019), and resection of any other gross visible disease (Armstrong et al., 2021). Comprehensive staging via multiple biopsies and lymph node dissection (according to the stage) usually follows primary surgery (Kurnit et al., 2021). Recent guidelines recommend that patients with early-stage disease undergo lymph node dissection (Kurnit et al., 2021). There was no difference in progression-free survival (time from treatment initiation until disease progression (Hess et al., 2019)) or overall survival (duration of patient survival from the time of treatment initiation (Hess et al., 2019)) among patients with advanced disease stage either with or without lymph node dissection in a recent randomized trial (involved 647 patients) (Harter et al., 2019). After the debulking surgery, according to their tumor histology type and substage, patients with early disease (stage I) have the option to remain without chemotherapy for observation (Armstrong et al., 2021) since chemotherapy following surgery did not show any advantage compared to undergoing surgery alone in these patients (Young et al., 1991; Lawrie et al., 2015).

Following the primary cytoreduction surgery, most patients (stages II-IV) generally receive systemic chemotherapy to reduce the risk of disease recurrence or to treat residual disease (Armstrong et al., 2021). For more than two decades, a combination approach incorporating platinum-based and taxane-based drugs remain the standard chemotherapy to treat primary ovarian cancer patients (Ozols et al., 2003). In the past, older patients with significant disease burdens were often treated with single-agent platinum-based chemotherapy (Kurnit et al., 2021). However, the current practice strongly recommends the combination chemotherapy based on the results of a randomized clinical trial that showed improved outcomes of patients who received the combination compared to those who received the single agent (Carboplatin) (Falandry et al., 2019).

In selecting the right platinum-based chemotherapeutic agent, multiple randomized trials compared Carboplatin to Cisplatin efficacy and toxicity (du Bois et al., 2003; Ozols et al., 2003; Hilpert et al., 2007). All trials indicated their equivalent efficacy; however, Cisplatin incorporation caused higher levels of non-specific toxicity than Carboplatin (du Bois et al., 2003; Ozols et al., 2003; Hilpert et al., 2007). Cyclophosphamide, Paclitaxel, Docetaxel, and Doxorubicin are examples of chemotherapeutic agents used in combination with platinum-based agents (Hakes et al., 1992; McGuire et al., 1996; Parmar et al., 1998). Paclitaxel demonstrated a better response rate, progression-free survival, and overall survival compared to the other agents (Armstrong et al., 2021), which placed the combination of Carboplatin and Paclitaxel in the position of "standard" combination first-line therapy option for postoperative ovarian cancer patients. Following the first-line

chemotherapy, patients undergo regular clinical evaluations to monitor the success of previous interventions (Armstrong et al., 2021). Patients with complete remission, partial remission or stable disease usually receive maintenance chemotherapy such as PARP inhibitors (PARPi) depending on their case-to-case evaluations (Armstrong et al., 2021).

The PARP family of enzymes consists of seventeen members that vary in structural domains, subcellular localization, and functions (Gupte et al., 2017). They use nicotinamide adenine dinucleotide (NAD+) as a substrate to add ADP-ribose (ADPr) units to acceptor proteins as post-translational modifications (ADP-ribosylation) (Bürkle, 2005). These ADPr units vary in length from shorter mono(ADP-ribose) (MAR) to longer poly(ADP-ribose) (PAR) (Vyas et al., 2013). Based on their functionally characterized domains in regions outside the PARP domain (their structural domains and functions), the PARP family can underlie four different subfamilies (Vyas et al., 2013). The first three members are DNA-dependent PARPs that contribute to DNA DSB repair by ADP-ribosylating histones. This PARylation of histones promotes chromatin accessibility to enable the recruitment and function of repair machinery (Gupte et al., 2017). Although PARP-2 and PARP-3 have a similar activation mechanism as PARP-1, they require 5' phosphorylation at the DNA breaks for activation, unlike PARP-1 (Gupte et al., 2017). Tankyrases (PARP-5a and PARP-5b) contain protein-binding ankyrin repeats that are implicated in protein-protein interactions to regulate cellular processes such as telomere maintenance, Wnt signaling, glucose metabolism, and mitosis (Smith and De Lange, 2000; Chang et al., 2005; Huang et al., 2009; Guo et al., 2012). The members of the third PARP subfamily (PARP-7, PARP-12, PARP-13.1, and PARP-13.2) have CCCH zinc

finger domains that can bind to viral RNA (Vyas et al., 2013) as previously mentioned in section 1.4.3. The macro-PARPs (PARP-9, PARP-14, and PARP-15) contain ADPr-binding macro domains capable of recognizing MARs and PARs and might play a role in supporting PARPs biological functions by providing a molecular link between distinct family members (Karras et al., 2005). PARPs are involved in stress responses, cellular metabolism, and immune functions (Ryu et al., 2015; Gupte et al., 2017; Kim et al., 2020).

Most available PARPi target a relatively broad spectrum of the PARP family. However, Olaparib shows increased selectivity for PARPs-1-4 (Wahlberg et al., 2012). BRCA1-deficient cancer cells are sensitive to PARP inhibition. They survive by relying on PARP-mediated DNA repair (Foulkes and Shuen, 2013), so inhibition of PARPs in these cells (with BRCA1 loss of function) leads to synthetic lethality and a consequent improvement in ovarian and breast cancer treatment efficacy (Farmer et al., 2005). This synthetic lethality occurs only when both repair pathways (BRCA1- mediated and PARP- mediated) are deficient in cells (Farmer et al., 2005).

Since its approval by the EMA (European Medicines Agency) and FDA (Food and Drug Administration) as a treatment for patients with BRCA1/2 mutated advanced-stage epithelial ovarian cancer, Olaparib has shown significant improvements in the progression-free survival rates during phase III clinical trials regardless of BRCA1/2 status (Moore et al., 2018; Banerjee and Lord, 2020).

1.6.4 Retrotransposons as attractive targets in ovarian cancer

Multiple lines of evidence suggested that epithelial ovarian cancer involves moderate retrotransposon activity (Lee et al., 2012; Helman et al., 2014; Rodić et al., 2014; Tang et al., 2017). Retrotransposon expression levels are elevated in malignant ovarian tumors, possibly due to their CpG hypomethylation (Menendez et al., 2004; Pisanic et al., 2019). Some evidence suggests that retrotransposon activation in ovarian cancer occurs gradually during tumor evolution and can affect the clinical outcomes (Nguyen et al., 2018). However, little is known about the causes of retrotransposon triggering and the mechanisms regulating its consequences in ovarian cancer. Retrotransposons can serve as attractive targets for improved ovarian cancer detection, leading to an earlier diagnosis and better overall survival. As a first step, it is imperative to understand the mechanisms controlling the retrotransposon activation in ovarian cancer. Therefore, in this study, we wanted to shed light on retrotransposon activity causes and consequences in ovarian cancer, specifically in the high-grade serous subtype. We used patient data deposited in public databases, retrotransposon consensus catalogues and available bioinformatics tools to analyze retrotransposon activity at a genome-wide scale. In addition, cell models were employed to validate our findings in patient datasets and understand some of the underlying regulatory mechanisms

2 CHAPTER 2: GENOME-WIDE STUDY OF RETROTRANSPOSONS IN OVARIAN AND BREAST CANCER

2.1 Background

A race is going on between numerous research groups to study and understand mechanisms governing retrotransposon activity in cancer. The driving reasons may include the substantial abundance of retrotransposons in the human genome, the nature of their activity, and their association to cancer that has been established over the last decades. Retrotransposons are involved and could contribute to multiple cancer hallmarks such as genomic instability, associated inflammatory responses, and telomere maintenance (Hanahan and Weinberg, 2011; Sudmant et al., 2015; Aschacher et al., 2016; Kong et al., 2019). They can also represent attractive targets for cancer therapeutic developments.

Recent advances in molecular biology techniques and bioinformatics tools have paved the way for studying retrotransposons from different perspectives at high throughput levels. However, because of their repetitive nature and sequence variabilities, they are often 'masked' with other repetitive elements in the genome to avoid false-positive results associated with the repeats and accelerate downstream computational analysis (Bedell et al., 2000). This exclusion of retrotransposons presents a challenge to add to other challenges in analyzing and studying them, such as the difficulty in precisely determining

their insertion boundaries. This difficulty can be related to the retrotransposon sequence characteristics or the quality of available data. L1 sequence, for example, differs among genomic copies in their polyadenylation signal, and 3' UTR with most copies being 5' truncated (Scott et al., 1987; Boissinot and Furano, 2001; Szak et al., 2002). Most available whole-genome sequencing (WGS) data consists of single or paired-end short reads of about 100-250 nt in length (Ewing, 2015). Using these reads to detect TE insertions in sequences containing a background noise of thousands of similar interspersed copies presents another challenge to resolve. Therefore, filters and measures to control for false-positive rearrangements are required to detect TE insertions with reasonable sensitivity (Ewing, 2015).

Large-scale sequencing projects that include data from thousands of patients deposited in public databases such as The Cancer Genome Atlas project (TCGA) and the International Cancer Genome Consortium (ICGC) and the development of bioinformatics tools and pipelines has facilitated the comprehensive detection and analysis of retrotransposons in cancer (Goerner-Potvin and Bourque, 2018). Available tools that accelerated research in the TE field and are relevant to our study can range from data repositories to insertion detecting tools and strategies to investigate their biological impacts. Databases such as RepBase Update and the European database of L1HS retrotransposon insertions (EUL1Db) were developed as repositories focused on assembling TE families consensus sequences and polymorphic non-reference TE insertions, respectively (Jurka et al., 2005; Mir et al., 2015).

Two elements are required to identify TE polymorphisms in an individual sequenced genome, available reference genome and annotated TE sequences. Both of which are made accessible in the public databases. Detection of TE polymorphisms is achieved by accurate alignment of reads to the reference genome, looking for deviations from the reference sequence (Ewing, 2015). Some of the identified polymorphic insertions were linked to diseases such as hemophilia (Kazazian et al., 1988) and Rett syndrome (Yu et al., 2001). Many TE detection software tools based on robust algorithms have been developed to identify germline and somatic TE insertions using sequencing short reads as in the TCGA (Goerner-Potvin and Bourque, 2018). Short reads do not span the entire interval affected by retrotransposon-mediated genomic rearrangement (Ewing, 2015). Therefore, computational tools were developed to utilize up to three strategies in detecting TE insertions: inference from discordant read pair (DRP) mapping, clustering of split-reads (SR), and sequence re-alignment through the identification of TE-specific motifs (Rishishwar et al., 2017).

DRP methods detect a pair of reads from the same TE insert whose alignments to the reference sequence (if contiguous on it) have orientation or distance that differ from the expected range (Elyanow et al., 2018). No identification of exact junctions between TEs and the reference genome is possible using DRP methods alone (Ewing, 2015). In contrast, SR methods detect reads with no contiguous alignment to the reference sequence, but they overlap partially with the surrounding genome and the TE sequences (Elyanow et al., 2018). In SR, a read segment aligns to the reference genome sequence, and other segments map to a part of the TE sequence (Elyanow et al., 2018). Nonreference SRs are clipped to align with the reference sequence and can be used to identify the junctions between the

TE and reference genome sequence (Ewing, 2015). Therefore, SR strategies provide a higher positional accuracy by identifying the junction between the TE and host sequence. DRP strategies, on the other hand, offer higher sensitivity providing more reads to support TE insertions (Goerner-Potvin and Bourque, 2018). However, another strategy is required to refine the DRP mapping and exclude TE-unrelated rearrangements such as SR or TE-specific motif detection (Ewing, 2015). In the TE-specific motif detection strategy, tools were developed to identify insertions by looking for common TE signatures such as target site duplications (TSDs) flanking most TE insertions, long stretches of poly(A) tails and 3' transduction in L1-mediated insertions (Goerner-Potvin and Bourque, 2018).

One of the recently developed tools for the 1000 Genomes Project that combined both SR and DRP approaches to offer accuracy and sensitivity in identifying and detecting TE insertions is The Mobile Element Locator Tool (MELT) (Gardner et al., 2017). MELT detects mobile element insertions (MEIs) by searching for signatures of DRPs and SRs, enriched at sites containing novel, non-reference MEIs in WGS data (Gardner et al., 2017). At a population scale, the application of MELT allows the construction of MEI models using available DRP and SR data from numerous samples to accurately locate each MEI site and identify its features (Gardner et al., 2017). A comprehensive set of MEI-associated features can be identified through MELT including, chromosomal insertion site, MEI orientation, TSD, internal mutation profile, and subfamily (Gardner et al., 2017). MELT provides a map of polymorphic MEIs in a given genome by genotyping samples for new (non-reference) and reference mobile element copies (Gardner et al., 2017). Furthermore, MELT can evaluate the potential impacts of MEIs on nearby genes

and identify impacted gene features such as promoter, coding exon, intron, UTR, or terminator (Gardner et al., 2017).

The TE field advances have been extended to offer tools to predict the impacts of TEs on gene regulation, such as measuring the overlap with other genomic regions, looking for associations to transcription regulation datasets or considering signs for negative or positive selection (Goerner-Potvin and Bourque, 2018). Looking for the active TEs studying the effect of these elements on the expression of nearby genes, the RepEnrich tool was developed (Criscione et al., 2014), which identifies TEs in Chromatin Immunoprecipitation (ChIP) sequencing and RNA sequencing data (Criscione et al., 2014). RepEnrich tool creates a series of contiguous segments representing all TE instances of each TE subfamily annotated in the TE repository (e.g., Repbase in this study) (Criscione et al., 2014). These series are then used to identify reads that map only to one subfamily of TEs, such as L1HS. Reads identified using this tool can be described as unique to a particular subfamily in the genome.

As summarized in Table 2-1, genome-wide studies followed multiple strategies in studying retrotransposon activity and discovered somatic L1 insertions in cancer. However, they all fall under one of two categories: targeted resequencing assays such as retrotransposon capturing sequencing and bioinformatics analysis of WGS or whole-exome sequencing (WES) data (Scott and Devine, 2017).

Table 2-1 Retrotransposon activity in cancer genome-wide studies

Citation	Data used (database)	Sample size	Strategy	Focus	Important Findings
(Lee et al., 2012)	WGS (TCGA)	43	TE analyzer (DRP reads)	Identifying novel insertions	194 somatic TE insertions in tumors, biased toward hypomethylated regions. Tumors of epithelial origin showed more pronounced L1 activity than brain and blood cancer types
(Shukla et al., 2013b)	generated data	19	RC-seq	Identifying novel insertions	L1-mediated mechanisms enabling tumorigenesis in hepatocellular carcinoma, identified insertions in MCC and ST18
(Tubio et al., 2014)	WGS (TCGA and ICGC)	244	TraFiC pipeline (DRP reads)	Insertion characteristics and impact	2756 L1 somatic insertions in tumors with colorectal and lung cancers being the most affected. Insertions exhibited hypomethylated promotors by tracking down their sources. L1 insertions demonstrated minimal to no effect on the course of tumorigenesis
(Helman et al., 2014)	WGS, WES (TCGA)	967	TranspoSeq (DRP and SR reads)	Identifying novel insertions	810 somatic retrotransposon insertions in epithelial cancers and many of them occurred in known cancer genes (WGS). 35 novel somatic retrotransposon insertions (by WES) including an insertion into an exon of the PTEN
(Scott et al., 2016)	WGS	11	MELT (DRP and SR reads)	Identifying novel insertions	hot L1 insertion in APC gene in colon cancer
(Nguyen et al., 2018)	generated data	30	RC-seq	Identifying novel insertions	88 tumor-specific L1 insertions in ovarian tumors, and one intronic insertion added a novel cis-enhancer to STC1 gene and promoted chemoresistance in cells bearing this mutation
(Schauer et al., 2018)	generated data	35 patients, 10 mice	RC-seq	Identifying novel insertions	First report of L1 activity in HCC murine tumors, identified 8 L1 tumor specific insertions in 25 patients with alcohol abuse and 3 L1 insertions in 10 intra-hepatic cholangiocarcinoma patients
(Jung et al., 2018)	WGS, RNA-seq (TCGA, EGA, dbGaP)	298	Modified TE analyzer (DRP reads)	Identifying novel insertions and impact	L1 activity positively associated with *TP53 mutation* L1 insertion in exon of MOV10 low L1 activity in tumors with high immune signature
(Sultana et al., 2019)	generated data	28	ATLAS-Seq	Characteristics of L1 integration	L1 shows a broad capacity for integration into all chromatin states compared to other mobile elements. L1 integration is influenced by the replication timing of target regions, distribution of new L1 insertions differs from those of pre-existing L1 elements
(Rodriguez-Martin et al., 2020)	WGS, RNA-seq (PCAWG)	2,954	TraFiC pipeline (DRP reads)	Impact of insertions on structural variation	19,166 somatically acquired retrotransposition events, affected 35% of samples. L1 induced somatic structural variation in esophageal adenocarcinoma, and the second most frequent in head-and-neck and colorectal cancers.

Most studies focused on identifying new insertions and characterizing their effect on tumor modulating genes; however, little is known about factors that mediate retrotransposon RNA expression or factors triggered by its activation. Also, evidence suggests that releasing the epigenetic repression on L1 and increasing its RNA expression

can be immunogenic in tumor and aging models (Guler et al., 2017; De Cecco et al., 2019). This increase in retrotransposon expression led to less chemoresistance and better survival rates in tested tumor models (Guler et al., 2017). Therefore, understanding mechanisms controlling retrotransposon expression can help us understand cancer evolution, biology, and immunogenicity and potentially improve future therapies.

The purpose of this thesis is to fill gaps in our knowledge about the actual impacts of retrotransposons effects in the most aggressive types of ovarian and breast cancers (high-grade serous and basal-like type, respectively). Selecting the right tools for the study needs is essential to yielding comprehensive answers to our questions about retrotransposons at the insertion and transcription levels. Compared to other MEI discovery algorithms such as RetroSeq, Mobster, Tangram, and TEM, the MELT tool analysed WGS data with a shorter runtime and identified MEIs with higher sensitivity and specificity (Gardner et al., 2017). Also, it included a filter to recognise full-length L1 elements insertions through their 3' transductions (Gardner et al., 2017). Attempts to survey genome-wide TE transcription were criticized for excluding reads that deviate from the consensus sequence or those aligned to more than a single repetitive element subfamily (Criscione et al., 2014). This exclusion caused a significant fraction loss in reads that could impact transcription (Criscione et al., 2014). RepEnrich pipeline was developed to overcome the aforementioned limitations in strategies used to study TE transcription (Criscione et al., 2014). Therefore, as a strategy, we selected MELT to identify retrotransposon insertions in ovarian cancer and RepEnrich to study retrotransposon RNA expression in ovarian and breast cancer, as illustrated in Figure 2-1 below.

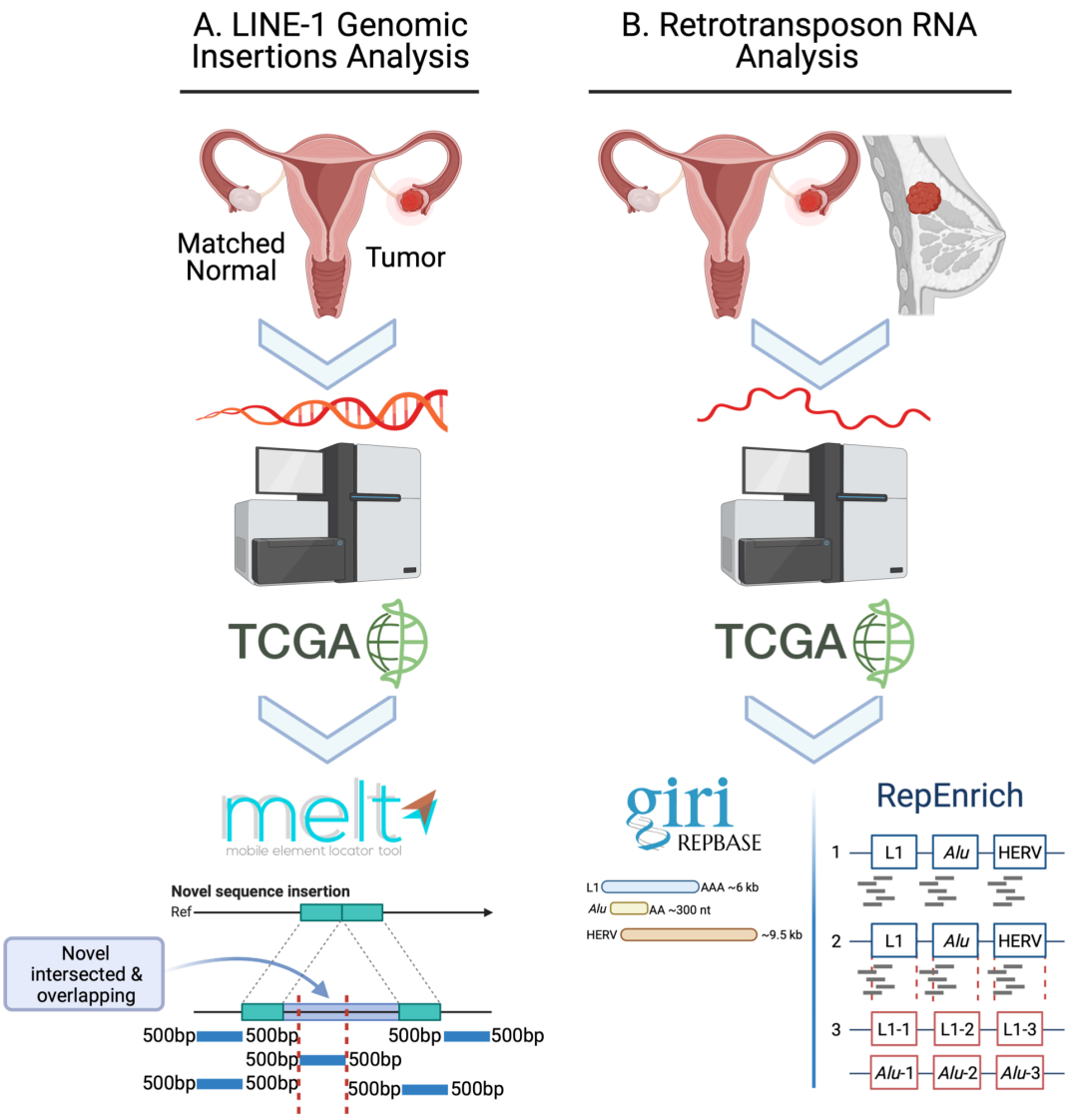


Figure 2-1 Genome-wide study workflow and methodology.

A. MELT (Gardner et al., 2017) was used to locate L1 genomic insertions across all chromosomes of 54 ovarian cancer patients by using their WGS data deposited in TCGA including tumor and their matched normal samples. 500bp on both sides of each insertion was used as a quality control to call for correct insertions within intersected and overlapping regions. **B.** RNA sequencing data from 379 ovarian cancer and 486 breast cancer patients was used to evaluate retrotransposon RNA expression. Active young element sequences included in Repbase (Bao et al., 2015) were quantified using RepEnrich (Criscione et al., 2014) by which reads are mapped to retrotransposons in the genome. A threshold is set to define regions with retrotransposon RNA reads across the genome. Only reads that map to the retrotransposon consensus sequences are aligned and then assigned to each retrotransposon subfamily. Illustrations were created with BioRender.

2.2 Results

MELT (Gardner et al., 2017) was used to evaluate retrotransposon insertions in tumors of patients with serous ovarian adenocarcinoma. It uses genome sequencing reads to map the retrotransposon insertion sites first by the DRP approach and then refines the junction position with the SR approach (Gardner et al., 2017). Genome sequencing data from all fifty-four ovarian cancer patients available in TCGA (including matched normal and tumor samples) was used in MELT to identify retrotransposon insertions with >20x coverage (Figure 2-2 A). Retrotransposon insertions that are catalogued in other genomes by EUL1Db (Mir et al., 2015) were considered likely inherited, and they were eliminated from the study along with insertions in genomic regions subject to frequent mapping errors. This elimination left 1,813 L1 insertions. Germline insertions specific to the patient or controlled populations found in both tumor and normal tissues numbered 1,121. These were also eliminated to leave 692 insertions that were considered tumor-specific. The analysed ovarian tumors presented a range of 0 to 64 *de novo* tumor-specific L1 insertions (Figure 2-2 B, Appendix A[A]). Although most tumors exhibited an evenly distributed range of tumor-specific (between 0- 30) L1 insertions, three patients had higher numbers of L1 insertions (40-64, Figure 2-2 B). Interestingly, normal tissue presented higher numbers of patient-specific L1 insertions than its tumor counterpart, and these numbers did not correlate with each other (Figure 2-2 C, Appendix A[A]). This evidence suggests that the number of tumor-specific L1 insertions in ovarian cancer patients is insignificant compared to the likely inherited or accumulated insertions over the patient's lifetime.

2.2.1 L1 Insertions Rarely Impact Tumor Suppressors or Oncogenes in Ovarian Cancer

In ovarian tumors, after excluding potential errors caused by repetitive sequences prone to mapping artifacts such as microsatellites within a sliding 500 bp window, we noticed L1 insertions broad distribution over chromosomes, but in clusters (Figure 2-2 D, Appendix A[B], B). A region encoding MHC class II HLA-DR contained the highest density cluster of novel L1 insertions. Some clustered insertions on both chromosome 20 (neighborhood of LINC01597) and chromosome 3 (neighborhood of EPHA3) were also observed, among others (Figure 2-2 E, Appendix A[B], B). L1 insertions in HLA-DR and EPHA3 regions have been associated with ovarian cancer growth (Kuusisto et al., 2013; Schuster et al., 2017). This evidence suggests that L1 has a tendency to insert in specific genomic regions over others, or inserting in these regions might provide a competitive advantage selected during tumor evolution.

Thirty insertions disrupted exons of protein-coding genes out of 692 L1 insertions in ovarian cancer genomes (Appendix A[C]), a nominal rate of gene disruption (30/692 insertions = 4.3% in exons, 1.1% of genome contains exons, $p>0.001$ Fisher's test). The putative oncogenes MMS22L, BCL11A and tumor suppressors ARHGEF12 and NEBL (CancerMine (Lever et al., 2019)) were interrupted by de novo L1 insertions. No enrichment was observed for L1 insertions in genes reported previously as essential for proliferation (Hart et al., 2015) (Figure 2-2 F). This observation generally implies that new L1 insertions in ovarian cancer are infrequently disrupting the coding regions of tumor suppressors, oncogenes or essential genes, and their interruption of tumor suppressors may rarely exert selective pressure in ovarian cancers.

Impacts of new tumor-specific L1 insertions can extend to change the status of DNA methylation or affect the expression of proximal genes by acting as promoters, enhancers or inhibitors. We observed no consistent effect of tumor-specific L1 insertions on DNA methylation or expression of proximal genes (within 10 kb, Figure 2-2 G, Appendix C). Although new L1 insertions were found in a small number of patients proximal to genes frequently mutated in cancer (e.g., BAZ1A, WT1) or to ovarian cancer tumor suppressor (VGLL3), the mRNA levels of only the latter gene was ranked in the lower percentiles (Appendix C). This evidence indicates that new L1 insertions only in rare cases affect the expression of proximal tumor-promoting genes.

Collectively, the above-shown evidence suggests that *de novo* L1 insertions minimally impact tumorigenesis or tumor progression in most ovarian cancer cases studied. Only three patients presented unusually high numbers of tumor-specific L1 insertions, and those died early to disease. Generally, in ovarian cancer, no correlation was observed between the number of L1 insertions and patient survival rates, as illustrated in Figure 2-2 H.

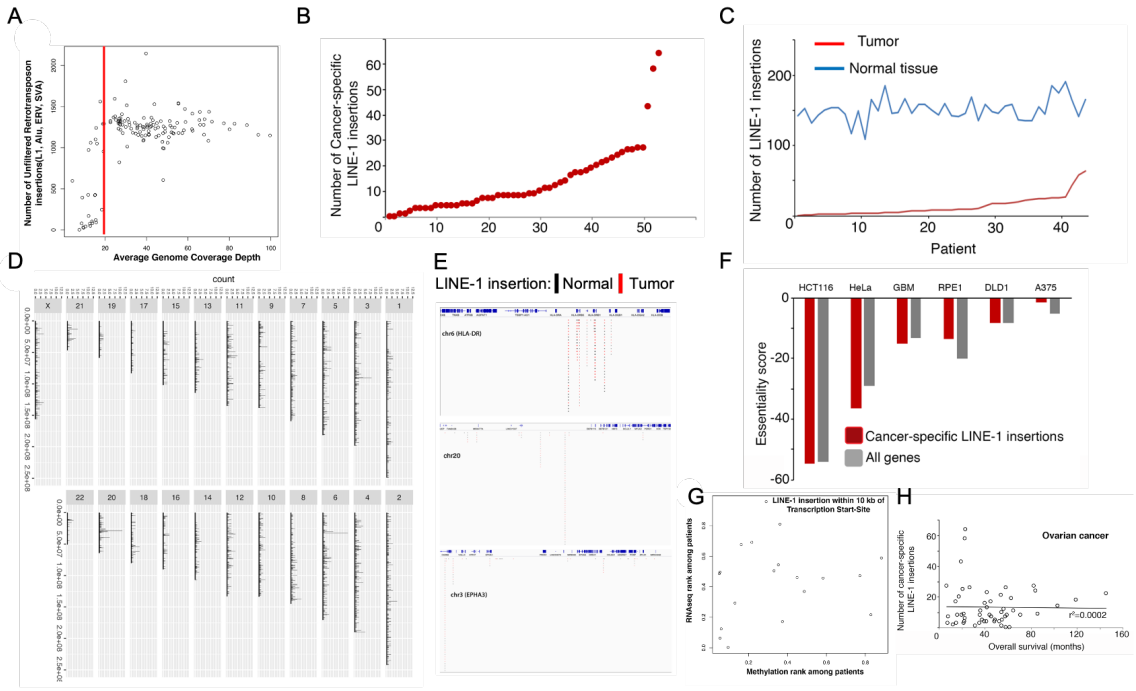


Figure 2-2 *De novo* L1 insertions in ovarian cancer.
A. Number plot of unfiltered retrotransposon insertions in the genome identified vs. the average genome coverage depth in the sample. Red line indicates average genome-depth of 20, where identification of retrotransposon insertions with MELT reaches an apparent plateau. **B.** Number of de novo tumor specific L1 insertions (y-axis) in each of 54 patients (x-axis). **C.** Number of de novo L1 insertions (y-axis) in either normal tissue samples or tumor samples from each patient (x-axis) **D.** Plot of de novo L1 insertions in tumor samples across 22 chromosomes. Peak height indicates count of L1 insertions. **E.** Examples of clusters of L1 insertions on chromosome 6, 20 and 3. Genes and exons are indicated in blue, while normal and tumor specific L1 insertions are denoted in black and red respectively. **F.** Essentiality score (Hart et al., 2015) for cancer specific L1 insertions vs. all genes in the indicated cell lines (x-axis). **G.** Plot of relative rank among all patients of RNA level vs. gene methylation level for the relevant gene with a de novo L1 insertion identified within 10 kb. **H.** Correlation plot of L1 genomic insertions vs. patient survival. Graphs were reproduced with permission from (Alkailani et al., 2021), Copyright Oxford University Press.

2.2.2 Retrotransposon RNA Predicts Survival of Ovarian, but not Breast Cancer Patients

Although genomic L1 insertions did not impact patient survival consistently, recent research demonstrated that chemically induced over-expression of retrotransposons and their RNA-DNA intermediates could be immunogenic and generate tumor repressive effects in animal models (Guler et al., 2017; De Cecco et al., 2019). We used serous ovarian adenocarcinoma and invasive breast carcinoma patient data from TCGA to determine whether retrotransposon RNA levels are sufficient, among tumor heterogeneity,

to impact patient survival significantly. RNA sequencing data from 379 and 486 patients were utilised, respectively. Repbase mapped retrotransposon RNAs (Bao et al., 2015) were quantified using Repenrich (Criscione et al., 2014) focusing on intact retrotransposition-competent RNAs including younger L1 (L1PA, and L1 HS), *AluY,* and HERVK families. We found L1 RNA levels to poorly correlate with levels of tumor-specific L1 insertions (Figure 2-3 A).

Intriguingly, ovarian cancer patients whose tumors expressed L1 in the highest quartile exhibited improved survival compared to those with L1 expression in the lowest quartile (Figure 2-3 B, C). Whereas data from breast cancer tumors analysed similarly showed no correlation between RNA expression and patient survival rates (Figure 2-3 D). The survival rates of patients could be affected by their tumor size, histological grade and metastasis status (Xiang et al., 2017). By analysing the characteristics of ovarian cancer patients included in the study, tumor grade, size, or invasion differences did not account for the observed difference in survival rates (Figure 2-3 E-H). This evidence suggests that factors stimulated by retrotransposons RNA or co-regulated with them improve the prognosis of ovarian, but not breast cancer patients.

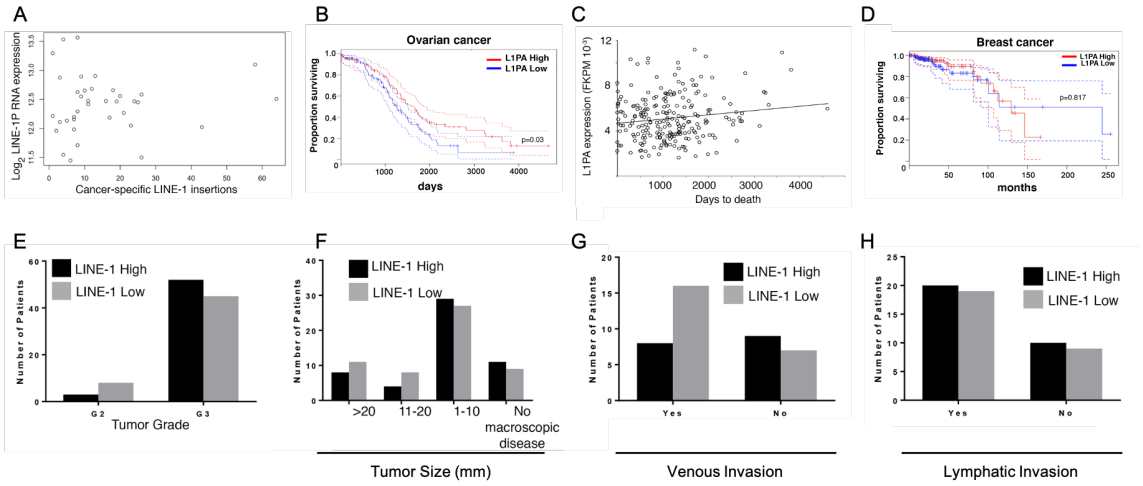


Figure 2-3 L1 RNA predicts survival of ovarian cancer patients.
A. Correlation of L1 RNA expression in tumors (counts per million) with number of tumor-specific de novo L1 insertions in ovarian cancer. **B.** Kaplan-Meier survival plots of ovarian cancer patients with top or bottom quartile expression of L1 PA RNA. Solid lines indicate average and broken lines indicate confidence interval. Error bars in all graphs represent standard error of the mean. **C.** Correlation of L1 PA expression (total proportion of mapped reads) and days to death of ovarian cancer patients. **D.** Kaplan-Meier survival plots of breast cancer patients with top or bottom quartile expression of L1 PA RNA. Solid lines indicate average and broken lines indicate confidence interval. Error bars in all graphs represent standard error of the mean. **E-H** Characteristics of cancer in patients whose tumors exhibit top or bottom quartile L1 RNA expression: (**E**) Tumor grade, (**F**) Tumor size, (**G**) Presence or absence of detectable venous invasion (**H**) Presence or absence of detectable lymphatic invasion. Graphs were reproduced with permission from (Alkailani et al., 2021), Copyright Oxford University Press.

2.3 Discussion

Retrotransposon activity in cancer has attracted much attention in the last decades, looking for potential therapeutic targets. With the ongoing genome-wide studies, regulation and impacts of retrotransposons are on the way to be revealed in different cancer types. Epithelial ovarian cancer is considered the most lethal gynecological cancer (Lheureux et al., 2019). Its advanced stage at the time of diagnosis contributes to the high death-to-incidence rate (Lheureux et al., 2019). Out of approximately 230 000 patients diagnosed with epithelial ovarian cancer, about 150 000 die annually worldwide (Lheureux et al., 2019). Established evidence of TE activity in epithelial tumors and the significant advancement in TE analysis tools suggest retrotransposons as attractive targets to investigate in the context of ovarian cancer.

Previous evidence suggested the involvement of retrotransposon activity in cancer initiation due to insertions in putative tumor suppressor genes disrupting their function or at proximity to oncogenes activating them (Morse et al., 1988; Miki et al., 1992; Shukla et al., 2013b; Scott et al., 2016). However, this was not the case in ovarian tumors data analysed in this study. Patient-specific L1 insertions number in normal tissue was much higher than in tumor tissue counterparts. Tumor-specific insertions number was insignificant compared to insertions likely inherited or accumulated throughout the patient's life. These insertions were infrequently disrupting the coding regions of tumor suppressors, oncogenes or essential genes. In rare cases, these rare insertion instances may have tumor promoting effects and impact the patient's survival, as in the three patients observed having unusually high numbers of L1 insertions and who succumbed early to disease. This observation agrees with previous reports indicating that retrotransposon insertions vary during tumor evolution and exhibit minimal effect on tumor progression (Lee et al., 2012; Tubio et al., 2014). Poor correlation of tumor-specific L1 insertions to L1 RNA expression might indicate that separate mechanisms govern the expression and insertion of L1 in different cancer types. This theory is supported by the observation that ovarian cancer patients with high L1 RNA expression survived longer, which was not apparent in breast cancer patients. In addition, the previous reports of immunogenic responses caused by overexpressing retrotransposons and the consequently favourable prognosis of cancer patients and animal models can support our interpretation further (Chiappinelli et al., 2015; Guler et al., 2017). Another article also demonstrated that better overall survival in ovarian cancer patients was associated with high HERV expression and consequent immunomodulatory effects (Natoli et al., 2021).

The developed bioinformatic analysis tools and the advances in retrotransposons studying systems facilitated the current discoveries in retrotransposons. Their activity is recognised in the cancer context, and extensive efforts are utilising available tools to understand the physiological impacts of this activity. For example, a study followed a comparable whole-genome and matched RNA-sequencing analysis of retrotransposon activity in gastrointestinal tumors but with different pipelines (Jung et al., 2018). As outlined in Table 2-1, Jung *et al.* used a modified TE analyzer (TEA) version that utilizes DRPs in mapping TE insertions. They improved the TEA tool to include a filter that selects insertions with both TSDs of at least 5 bp and poly(A) tails (Jung et al., 2018). In line with our observations, they found occasional instances of potential tumor-specific insertions in cancer genes, such as L1 insertions in tumors with *TP53* mutations and an insertion which interrupted *MOV10* tumor suppressor (Jung et al., 2018). However, unlike our findings, the tumor grade and patient age were positively correlated with L1 activity in the analysed tumors (Jung et al., 2018). This difference could be related to tumor-type-specific characteristics.

In general, we found tumor specific retrotransposon insertions to insignificantly impact ovarian cancer evolution. High L1 RNA expression predicted a significantly improved prognosis of ovarian cancer patients. An observation suggests studying retrotransposon RNA in the context of ovarian cancer to understand the mechanism behind this effect. The following chapter will discuss some of the analyses we carried out to expand our understanding of the causes and consequences of retrotransposons expression in ovarian cancer.

2.4 Methods

2.4.1 Retrotransposon Insertions, RNA Levels, and Methylation

Data from several sources (outlined below) were integrated based on the TCGA barcode patient and sample identifiers.

2.4.2 Transposon expression analysis

Aligned RNASeq TCGA (Weinstein et al., 2013) data for Ovarian and Breast Cancer patients were downloaded from the NCI Genomic Data Commons (GDC) (Grossman et al., 2016) made available through dbgap phs000178.v10.p8. This RNAseq data was generated by the TCGA from unstranded libraries using poly(A)-selected mRNAs. The raw read data reprocessed using the Repenrich analysis pipeline (Criscione et al., 2014) to generate read counts of transposons expression for *Alu,* L1HS, L1PA, L1PB and HERV. Expression of retrotransposon RNA was quantified as a proportion of total mapped reads or counts per million as indicated. These counts were integrated with TCGA provided HTSeq mRNA read counts.

2.4.3 Retrotransposon insertion detection

Paired-end WGS data for TCGA Ovarian cancer patients were downloaded from the NCI GDC (Grossman et al., 2016) made available through dbgap phs000178.v10.p8. Reads were analyzed using MELT (Gardner et al., 2017). Variable evidence in the form of per sample variation in read positions, depth, and insert size can lead to variations in MELT call positions for the same underlying insertion may vary by up to 500bp according to the MELT publication where they consider a MELT call to be correct if it is within 500bp of the known insertion. Each insertion position of cancer and normal samples from patients

was extended by 500bp on each side to a total width of 1kb. The intersection of these expanded insertions across all samples results in a set of overlapping regions. These regions represent the underlying insertion across samples. An intersection of the original insertion positions and the underlying insertion regions is assigned to a union of region overlapping insertions. By doing this, all called insertions are merged into a de-duplicated set. Within this set, insertions occurred in the tumor, but not normal samples of patients were considered potential cancer-specific insertions. We had a particular interest in these novel cancer insertions.

2.4.4 Methylation analysis

TCGA data is provided at four different levels defined by processing and accessibility terms (Silva et al., 2016). Level 1 refers to raw and controlled data. Level 2 refers to processed and controlled data. Level 3 indicates segmented or interpreted and open access data. Level 4 refers to a region of interest and open access data (Silva et al., 2016). In this chapter, TCGA level 3 data were used as sources for methylation values. These are derived from the most comprehensive platform for analyzing methylation data, Affymetrix human methylation 27k microarrays (Lister et al., 2009; Baker, 2010).

3 CHAPTER 3: RETROTRANSPOSON RELATED TYPE I INTERFERON RESPONSE IN OVARIAN AND BREAST TUMORS

3.1 Background

High-throughput sequencing technology can be considered one of the main drivers of recent advances in molecular biology research. Its flexibility enables measuring different types of molecules in different states such as DNA, mRNA, microRNA, DNA methylation, histone modifications and transcription factors binding (Sánchez-Taltavull et al., 2016). However, the precision of sequencing-based technologies differs among experiments and types of molecules being measured. Some of the factors which might influence the precision of sequencing data include experimental conditions, variance and the absolute level of the molecule being measured; for example, deeper sequencing results in greater precision (Sánchez-Taltavull et al., 2016). Also, high-count molecules are more precisely measured with a better signal-to-noise ratio than low-count molecules (Sánchez-Taltavull et al., 2016).

Linear studies usually include two continuous variables (e.g., x values and y values). the relation between x and y can be assessed by using Pearson correlation coefficient R. That represents a measure of association strength between x and y (ranges from -1 to 1) and can be defined by the following mathematical expression:

covariance (x, y)/ [standard deviation (x) X standard deviation (y)].

Every x value has a predictable y value on a best-fit regression line, and the better the association between x and y, the better one of them determines the other (Cleophas and Zwinderman, 2018). However, under certain conditions, Pearson and similar correlation methods that assume normal distribution can be limited by yielding high artificial similarities (Sánchez-Taltavull et al., 2016). Also, they can sometimes fall short when estimating accurate signal levels in case of uncertainty due to variances in signal levels between measured molecules and due to different sequencing depths between experiments (Sánchez-Taltavull et al., 2016).

The Bayesian correlation was developed to assess the similarity between different entities measured by high-throughput sequencing across different conditions (Sánchez-Taltavull et al., 2016). It measures the Bayes factor (BF) magnitude that is computed as the ratio of two likelihood distributions: posterior and prior distributions (Cleophas and Zwinderman, 2018). The prior distribution is based on the previously known background, whereas the posterior distribution considers the new information obtained (Cleophas and Zwinderman, 2018). In biological application, the prior distribution can integrate biological functional information into the model as prior knowledge and the posterior distribution is based on the measured parameters of interest (Cleophas and Zwinderman, 2018; Sauta et al., 2020). BF computation requires integrating prior and posterior likelihoods to maintain accuracy (Cleophas and Zwinderman, 2018). This integration makes Bayesian correlation advantageous over other analytical methods by considering the biological likelihoods when answering biological-related questions (Cleophas and Zwinderman, 2018).

Despite its potential in looking for accurate biologically relevant sequencing correlations, to our knowledge, Bayesian correlations were not applied to study retrotransposon expression or insertion in cancer-related research. The previously reported retrotransposon insertions in tumor-related genes and other accumulated information about retrotransposon activity in cancer would be helpful if integrated into the current genome-wide studies. Therefore, Bayesian correlation represents a potential tool to overcome challenges related to retrotransposons diversity and cancer heterogeneity, building on what is known already.

Availability of data is vital, but how researchers analyze it is more substantial to extract critical answers closer to the truth. For example, *de novo* retrotransposon insertions were estimated to be very high by a group of researchers (13.7 and 16.3 per neuron in the hippocampus and cerebral cortex, respectively) (Upton et al., 2015). Reanalysis of this data carried out by another group discovered experimental artifacts that raised the study's apparent rate of somatic retrotransposition by more than 50-folds (Evrony et al., 2016). The discovered artifacts were related to the sequencing approach, bioinformatic analysis, and validation methods (Evrony et al., 2016). Therefore selecting stringent bioinformatics tools that fit the study purpose and consider what has been done previously is of high significance to extract comprehensive answers for the raised questions.

This study sought to find more accurate answers by applying Bayesian correlations to survey factors regulating or regulated by retrotransposons in cancer. As demonstrated in the last chapter, high L1 expression predicted better survival rates of ovarian cancer, but not breast cancer patients. Our observations suggest the involvement of factors and

pathways present in ovarian cancer but absent in breast cancer. As mentioned earlier, evidence revealed that activating retrotransposons can be immunogenic and instigate IFN and apoptosis signaling (Roulois et al., 2015; Guler et al., 2017; De Cecco et al., 2019). Among the mechanisms that activate retrotransposons, demethylating agents such as DNMTi act by releasing the epigenetic restriction placed on retrotransposons (Roulois et al., 2015; Saito et al., 2016; Kong et al., 2019). Activation of various TE classes in glioblastoma cells triggered processing and presentation of TE-derived peptides on MHC class I molecules and consequent interleukin 1β or type I IFN responses (Kong et al., 2019). Activation of HERVs resulted in a viral mimicry response of dsRNAs inducing the MDA5/MAVS RNA recognition pathway and downstream activation of IRF7 (Roulois et al., 2015). Recent evidence (based on TCGA data analysis and in vitro DNMTi treatment of ovarian cancer cells) suggested that high HERV expression in patients was associated with better survival and correlated with infiltration of cytotoxic T cells (Natoli et al., 2021). However, little is known about L1 RNA expression consequences in ovarian cancer and whether it mimics viral response observed in cases of HERV. Our focus throughout this chapter will be on studying epigenetic regulation and inflammatory responses with retrotransposon RNA expression in ovarian and breast cancer by applying the Bayesian correlation method.

3.2 Results

3.2.1 A strategy to explore causes and consequences of retrotransposon RNA expression in ovarian and breast cancer

We aimed to broadly investigate mechanisms regulating retrotransposon RNA and the impacts of their expression in ovarian and breast cancers. As mentioned in the first chapter, retrotransposon expression can be controlled at the transcriptional level by heterochromatin and DNA methylation (Roulois et al., 2015; Guler et al., 2017); however, we observed no significant correlation between global DNA methylation and retrotransposon RNA levels (Figure 3-1 A). To identify RNAs whose expression correlated with each retrotransposon class, we used Bayesian correlation (L1, *AluY*, Figure 3-1 B and Appendix D[A]) in which high correlations are retained when measurement confidence is high and vice versa (Sánchez-Taltavull et al., 2016). Retrotransposition-competent sequences that match young TE elements were included in this analysis; however, the effect of retrotransposons embedded within other transcripts or other chimeric RNAs containing retrotransposon sequences could not be excluded. Other validating analyses are necessary to this study since the list of correlated RNAs will include a proportion of encoding proteins that directly control retrotransposon expression or are induced by them among many indirectly co-regulated RNAs with retrotransposons.

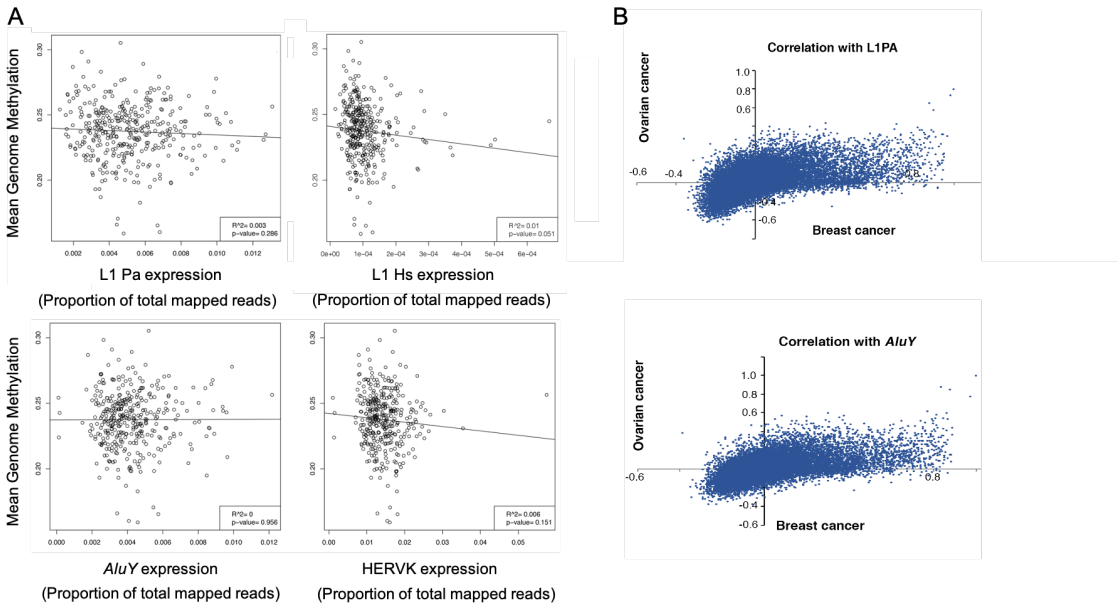


Figure 3-1 Retrotransposon expression correlations with DNA methylation and other RNAs
A. Correlation of L1Pa, L1HS (upper), *AluY* and HERVK (lower) RNA expression (Fragments Per Kilobase of transcript per Million mapped reads [FPKM]) with mean genome methylation in ovarian cancer. **B.** Plot of correlation values of each gene with L1Pa (upper) or *AluY* (lower) expression in ovarian cancer vs. in breast cancer. Graphs were reproduced with permission from (Alkailani et al., 2021), Copyright Oxford University Press.

3.2.2 Type I interferon responses are induced by and induce retrotransposon RNA expression

As a first assessment, we started with RNAs correlated with retrotransposon RNA levels and were previously reported to be stimulated by retrotransposons in cancer. In tumor and senescence models, retrotransposons were demonstrated to induce type I IFNs and promote apoptosis and immunogenic responses (Roulois et al., 2015; Guler et al., 2017). The evidence from Jung *et al*. that studied gastrointestinal tumors (discussed in chapter 2) described correlations of retrotransposon activity with the expression of immunoregulatory molecules, Toll-like receptors, and some mRNAs whose transcription is induced by IFN (Jung et al., 2018). In general, we observed these mRNAs having weak associations with retrotransposon expression in breast and ovarian cancer ($r < 0.2$, Appendix D[C]). Type I IFNs, including many paralogous IFNα mRNAs (IFNA21,

IFNA6, IFNA13, IFNA5, IFNA16), were among RNAs highly correlated with retrotransposon RNAs in breast cancer ($r > 0.4$, Figure 3-2A and Appendix D[A]). Particularly, type I interferon IFNκ was correlated at $r > 0.87$ with retrotransposon RNA in breast cancer. On the contrary, in ovarian cancer, all type I IFNs except for IFNε and IFNA5 ($r < 0.27$) presented very poor correlations with retrotransposon RNA ($r < 0.1$, Figure 3-2A).

We expressed L1 RNA in ovarian and breast cancer cell lines (ES2 and MDA-MB-231) to examine whether retrotransposon expression induces type I IFN. With a modest increase in L1 RNA, IFNκ mRNA expression was enhanced in both cell lines (Figure 3-2B). When type I IFN responses are induced by transfecting cells with poly I:C, L1 and *AluY* RNA expressions were also triggered intriguingly (Figure 3-2C). This observation suggests that L1 expression induces type I IFN as a response that stimulates further L1 expression in a self-amplifying loop evident in breast cancer but attenuated in ovarian cancer. Despite the high levels of type I IFN in breast cancer tumors accompanying high levels of L1 RNA (Figure 3-2A), no impact of L1 expression was observed on breast cancer patients' survival (Figure 2-3D), which suggests that high levels of type I IFN likely induced by retrotransposon RNA in breast cancer were insufficient to impact patient survival. This evidence also suggests that the effect of the L1 expression on the survival of ovarian cancer patients is not related to type I IFN responses.

Retrotransposon restriction factors such as TREX1, APOBEC3, and SAMHD1 had variant correlations with retrotransposon RNA. TREX1 and most members of the

APOBEC3 family had negative correlations with retrotransposon RNA in both ovarian and breast tumors as expected (Appendix D[A]). While APOBEC3A, APOBEC3D, and SAMHD1 were negatively correlated with retrotransposon RNA in breast cancer but not in ovarian cancer (Appendix D[A]). This observation suggests that the change in retrotransposon restriction mechanisms may contribute to the distinct phenotype noticed between the two cancer types.

Retrotransposons were previously shown to activate broad immune responses and recruit different immune cells into the tumor (Jung et al., 2018). The analysis of Cell-type-specific RNAs can serve as an indicator of specific cell types abundance in whole tissue (Wang et al., 2019). Applying this method facilitated the identification of cytokines and immune cells infiltrating tumors with high retrotransposon RNA expression. A negative correlation was observed between markers of cytotoxic T cells, monocytes and antigen-presenting cells and retrotransposon RNA in ovarian cancer that was less noticed in breast cancer (Figure 3-2D). Interestingly, IL-25 and its Th2 cytokine stimulants IL-4, IL-5 and IL-13 (Fort et al., 2001) exhibited the strongest positive correlations with retrotransposon RNA expression among analyzed cytokines in both breast and ovarian cancers (Figure 3-D).

Additionally, no association was noticed between Retrotransposon RNAs and inflammatory markers of the secretory-associated senescent response (Figure 3-2D), linked to retrotransposon activation in aging (De Cecco et al., 2019), suggesting that retrotransposon RNA is mainly associated with a Th2 and Th9-type immune response. Although the cause of this is unclear, it could be inferred that tumor cells with abundant

retrotransposon RNA develop a competency to avoid killing by enabling the Th2 environment to repress cytotoxic responses consequently.

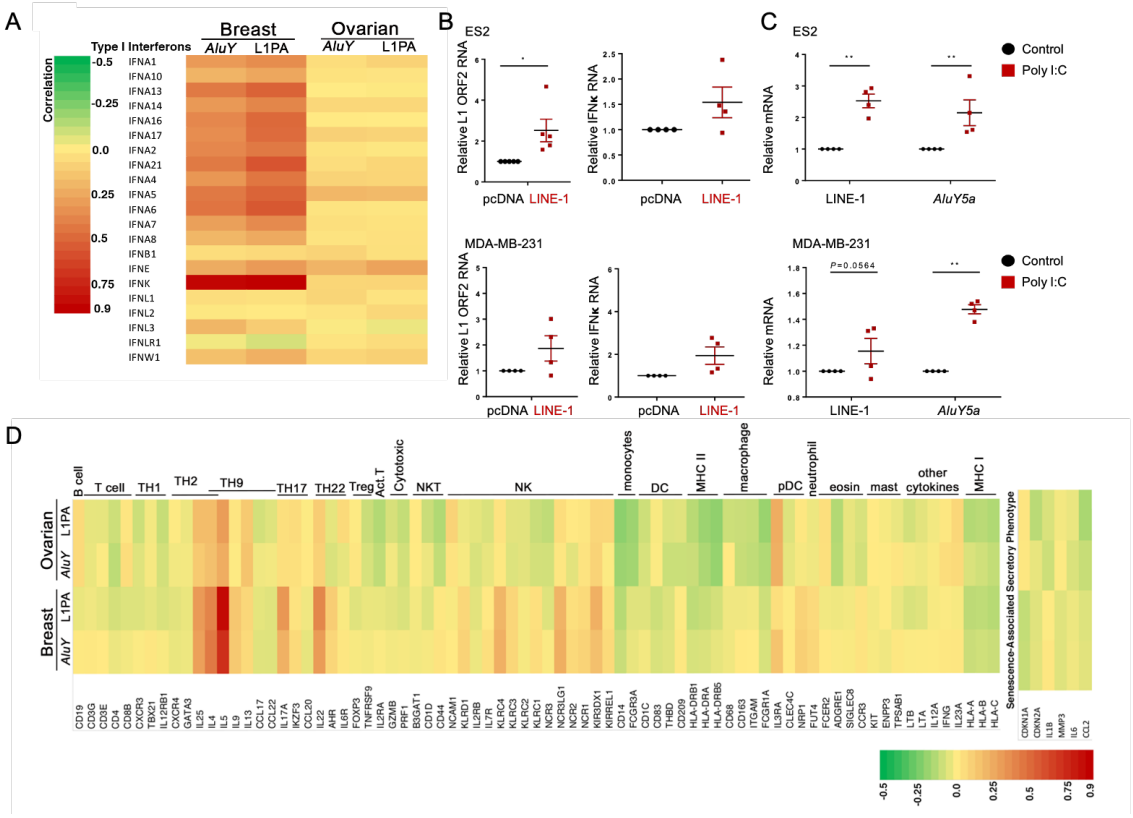


Figure 3-2 Retrotransposon Expression Correlations with Inflammatory Mediators
A. Heat map of correlation values of type I IFNs with L1 or *AluY* RNA in ovarian and breast cancer. **B.** RT-qPCR of L1 ORF2 (left) or IFNκ (right) mRNA after transfection of ES2 (upper) or MDA-MB-231 (lower) cells with L1 expressing plasmid vs. control plasmid. N=4-5 independent biological replicates. **C.** RT-qPCR of L1 ORF2 or *AluY* RNA after transfection of ES2 (upper) or MDA-MB-231 (lower) cells with the activator of anti-viral responses poly I:C (1µ g/ml). N=4 independent biological replicates. **D.** Heat map of correlation values of immunological responses or the senescence associated secretory phenotype with L1 or *AluY* RNA in ovarian and breast cancer. In RT-qPCR horizontal lines represent averages and error bars indicate standard error of the mean. *p<0.05, **p<0.01, two-way ANOVA with Holmes-Sidak correction (C) t-test (B). Graphs were reproduced with permission from (Alkailani et al., 2021), Copyright Oxford University Press.

3.3 Discussion

Literature on retrotransposons suggests that DNA methylation and IFN signaling are involved in regulating retrotransposons activity and the consequences of their activation, respectively. Although L1 activity in cancer is strongly associated with general

hypomethylation status (Hur et al., 2014; Tubio et al., 2014), we observed no significant effect of global DNA methylation on retrotransposon RNA expression in ovarian tumors (Figure 3-1 A). This evidence suggests that improved patient survival observed with high L1 RNA expression is independent of DNA methylation status. The favoured effect of demethylation drugs noticed in ovarian cancer survival rates previously was not entirely related to the genome methylation state but to HERV expression and its consequent responses (Natoli et al., 2021).

Immunoregulatory molecules such as Toll-like receptors and mRNAs whose transcription is induced by IFN showed weak associations to retrotransposon expression in breast and ovarian cancer, unlike recent findings in gastrointestinal tumors (Jung et al., 2018). Tumor type and microenvironment may play a role in retrotransposon expression and the observed differences. For example, IFNε , which presented one of the highest correlations with retrotransposon expression in breast and ovarian cancers, is hormonally regulated and expressed in cells of reproductive organs (Fung et al., 2013). Although estrogen hormone treatment activates one of the central repressors of retrotransposons (APOBEC3) (Pauklin et al., 2009), the 5'UTR of HERVK contains multiple binding motifs for estrogen and progesterone hormones, and they were shown to upregulate its expression (Mustelin and Ukadike, 2020). Evidence shows that retrotransposons contain several binding motifs for estrogen response elements (ERE) (Mason et al., 2010). Therefore, the effect of hormones on retrotransposons could be related to a direct binding between ERE elements and retrotransposon DNA.

One of the ERE is NQO1, quinone reductase (NAD(P)H:(quinone acceptor) oxidoreductase, which is activated by blocking estrogen binding to its receptor (Montano et al., 1998). Interestingly NQO1 expression is negatively correlated with retrotransposon expression along with other EREs such as EBAG9, LTF, CTSD, GATA3 (Ikeda et al., 2015) (Appendix C). This observation suggests that a mechanism specific to hormone-responsive tissue may regulate retrotransposon expression in ovarian cancer. Also, with a less clear mechanism, the use of hormonal contraceptives was associated with an increased risk of developing breast cancer and a decreased risk of developing ovarian cancer (Cook et al., 2017; Mørch et al., 2017). The involvement of tissue specificity in the regulation of retrotransposons is an exciting topic open for further research.

The tumor microenvironment is occupied by various immune cells, including members of innate and adaptive immune lineages. Cytotoxic T lymphocytes are recruited for protection against tumor development in acute tumor-directed immune responses. In contrast, the infiltration of Th2 cells is involved in chronic activation of humoral immunity that promotes tumor development and disease progression (Setrerrahmane and Xu, 2017). We noticed a general negative correlation between retrotransposon expression and markers of cytotoxic T cells, monocytes and antigen-presenting cells in ovarian cancer. Th2 signatures were the most correlated in breast cancer and to a lesser extent in ovarian cancer. These cytokines (IL-25, IL-4, IL-13) were shown previously to promote ovarian and breast cancer cell proliferation, and invasion and their inhibition benefited tumors prognosis (Kioi et al., 2005; Furuta et al., 2011; Fujisawa et al., 2012; Cao et al., 2016; Gaggianesi et al., 2017; Jiang et al., 2017). In addition, our observations of L1 expression inducing type I IFN as a response that stimulated further L1 expression all to indicate that

these responses might be associated with disease progression due to chronic activation of Th2 cells. However, our data also indicates attenuation of this response has occurred in ovarian cancer since that inflammatory response was more prominently correlated with retrotransposon RNA in breast cancer. Even with this, the high correlated inflammatory responses were not enough to affect patients' survival in breast cancer.

This study was limited by not including controls such as inhibitors of DNA methylation or RTi in the experiments and missed validating inflammatory markers other than IFNκ. Including these in future research will add support to the findings and may open different aspects to investigate. Nevertheless, our observations from Bayesian correlation analysis, besides our in vitro experiments in cell models, suggest no association between DNA methylation and inflammatory responses and the demonstrated effect of retrotransposon expression on improved survival of ovarian cancer patients. Retrotransposon levels could be insufficient to trigger inflammatory responses in ovarian cancer. Another possibility could be related to a higher threshold to stimulate these responses in ovarian cancer than other cancer types. Also, a complex network of genes correlated with retrotransposon RNA might predict the survival of ovarian cancer patients in new mechanisms other than DNA state or inflammatory response, which will be discussed further in the next chapter.

3.4 Methods

3.4.1 Transposon expression analysis

As described in section 2.4.1 with the addition of Bayesian Pearson correlation (Sánchez-Taltavull et al., 2016) of retrotransposon versus the gene expression counts was performed (see Appendix D).

3.4.2 Methylation analysis

As described in section 2.4.3, with the addition of Figure 3-1A information that represented the mean methylation of all probes (an estimate of genomic methylation) versus L1PA CPM (Counts per million mapped reads), as determined using Repenrich (Criscione et al., 2014).

3.4.3 RT-qPCR

RT-qPCR was performed after RNA isolation from cells using TRIzol (Invitrogen). RNA samples were DNaseI (Qiagen, 79254) treated, and cDNA was prepared using MiScript II Reverse Transcriptase (Qiagen). GoTaq qPCR Master Mix (Promega A6002) was used for qPCR with primers as listed in Table 3-1 below. Relative quantities were calculated using the ΔΔ*C*t method, normalized to the geometric mean of β-actin and GAPDH unless otherwise indicated.

Table 3-1 Oligos used in chapter 3 sequences

Target	Forward Oligos (5'-3')	Reverse Oligos (5'-3')
L1-ORF2	TGC GGA GAA ATA GGA ACA CTT TT	TGA GGA ATC GCC ACA CTG ACT
Alu-Ya5	TCC CGG CTA AAA CGG TGA AA-3	CTC CCA AGT AGC TGG GAC TAC AGG
GAPDH	ATC TTC TTG TGC AGT GCC AG	TTT GCC ACT GCA AAT GGC AG
β-actin	GGC ATG GGT CAG AAG GAT T	GGG GTG TTG AAG GTC TCA AA
IFNκ	GCC CCA AGA GTT TCT GCA ATA C	GGC CTG TAG GGA CAT TTC ATA GA

3.4.4 Cells and Reagents

MDA-MB-231 cells (from Dr. Jocelyn Côté) were cultured in Dulbecco's Modified Eagle's Medium containing 10% fetal bovine serum and 2 mM L-glutamine. ES2 cells (from Dr. Barbara Vanderhyden) were cultured in Modified McCoy's 5a Medium with 10% fetal bovine serum. Lipofectamine 2000 (Invitrogen) was used to transfect Poly I: C (R&D, 4287/10) or the following plasmids into cells: pCEP 5'UTR ORF2 no-Neo, *Alu*-neo^{Tet}, pcDNA3.1, L1-neo-TET (Kroutter et al., 2009) (Addgene # 51284).

3.4.5 Statistical Analysis

Two-tailed t-test and two-way ANOVA were employed to evaluate the statistical significance of experiments as appropriate in GraphPad Prism software on a minimum of three independent biological replicates. A *P*-value of <0.05 was considered statistically significant. Significance was denoted as follows: * *P*-value < 0.05, ** *P*-value < 0.01, *** *P*-value < 0.001, **** *P*-value < 0.0001

4 CHAPTER 4: CANDIDATE MECHANISMS AND GENES REGULATE RETROTRANSPOSON RNA IN OVARIAN AND BREAST CANCER

4.1 Background

The fast-growing knowledge of molecular sequences, including the entire genomes of model organisms, has transformed experimental biology's theory and practice (Ashburner et al., 2000). This advancement contributed to an understanding of shared genes and proteins among diverse organisms and the knowledge of their biological roles (Ashburner et al., 2000). Each protein can carry out different molecular functions in several processes and contribute to alternative interactions with other proteins (Ashburner et al., 2000).

The biological community needed to find a common language to communicate agreed-on terms of all genes and gene-products to indicate their functions and interactions along the discovery line. The Gene Ontology (GO) resource was created in 1998 to support this cause, and since then, it is under continual improvement, both in quantity and quality, as more information accumulates (The Gene Ontology, 2019). As mentioned earlier, many valuable tools were developed to facilitate the understanding of retrotransposon biology, regulation, and impact on cellular functions. There is much yet to uncover about networks of RNAs or mechanisms that regulate retrotransposon RNA expression in cancer.

The type I IFN analysis demonstrated in the previous chapter indicates that Bayesian correlation has predictive power in identifying RNAs controlling retrotransposon expression or induced by them. This chapter will explore other candidate genes and processes, beginning with an analysis of associated GO terms for possible regulation of retrotransposon expression.

4.2 Results

4.2.1 Retrotransposon RNAs correlate to each other, and transcription factors regulating their expression

Looking to identify new factors using the Bayesian correlation approach, we started with correlations between different classes of retrotransposons. Levels of L1, *AluY*, and HERVK RNAs exhibited remarkably high positive correlations with each other in all combinations (r=0.55 to r=0.93) than other mRNAs in ovarian and breast cancer (Figure 4-1 A,B). By incorporating RNA derived from L1HS, the more active L1 family in the ovarian cancer analysis, we also observed a close correlation with other classes of retrotransposons, including the broader L1PA family (Figure 4-1 A). Generally, both L1 families correlated similarly with mRNAs in the analysis (Appendix D[B]). In line with this, *AluY* RNA levels were augmented upon expressing L1 in cells (Figure 4-1 C), potentially due to the stabilization of *AluY* RNA when it is bound to L1 ORF1p. Some tumor properties such as those controlling transcription, epigenetics or RNA stability might act on pathways affecting all retrotransposon classes and accordingly contribute to the tight correlation among retrotransposon classes in these cancers.

For example, we observed genes correlated with L1 expression in ovarian cancer to be highly enriched in binding sites for the YY1 and Oct1 transcription factors (Appendix D[D]). Similarly, genes co-regulated with retrotransposon RNA in breast cancer were enriched in binding sites for Oct/Pou family, MEF2C family, and SATB1(Appendix D[D]). L1 expression was previously shown to be controlled by YY1 binding sites in the 5'UTR promoter of L1 element (Athanikar et al., 2004), Oct2 was demonstrated to bind L1 elements (Sun et al., 2018) and SATB1 was found to bind *Alu* sites in the genome (Kumar et al., 2007). This evidence suggests that these transcription factors may control a gene expression program that includes retrotransposon stimulation in ovarian and breast cancer.

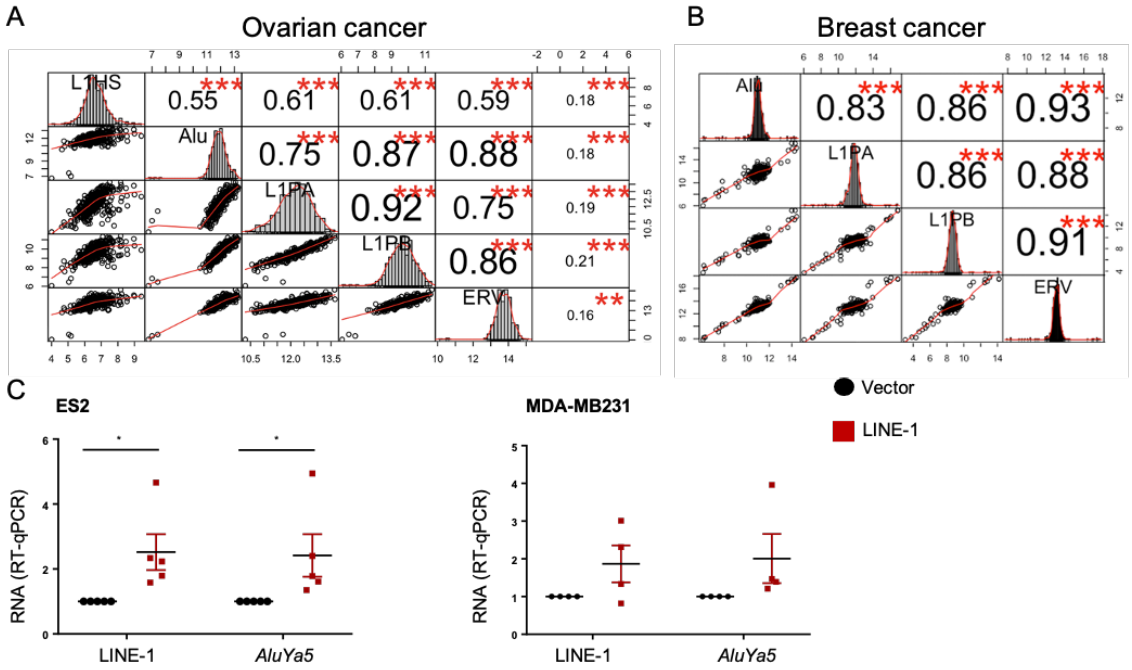


Figure 4-1 Retrotransposon RNAs are closely correlated to each other in both ovarian and breast cancers.
A,B. Pearson correlation (r) of L1 families, *AluY*, and HERVK RNA (proportion of total reads mapped) in ovarian cancer (**A**) and breast cancer (**B**). Dot plots represent correlation of mean expression values per patient. Histograms represent distribution of expression of indicated retrotransposons. Numbers indicate the r correlation value. **C**. RT-qPCR of L1 (ORF2) and *AluYa5* retrotransposon RNA two days after transfection with plasmid expressing L1. N=4-5 independent biological replicates. Data on L1 RNA levels is also used in Figure 3-2B. In RT-qPCR horizontal lines represent averages and error bars in all graphs represent standard error of the mean. *p<0.05 two-way ANOVA with Holmes-Sidak correction. Graphs were reproduced with permission from (Alkailani et al., 2021), Copyright Oxford University Press.

4.2.2 A network predictive for L1 expression identifies regulation by mitochondrial activity

To identify additional candidate processes or complexes that govern retrotransposon RNA expression, we assessed the interactome and GO-terms (Ashburner et al., 2000; Carbon et al., 2019) of RNAs highly correlated with retrotransposon expression. A concentration of histones and G protein-coupled receptors was detected by GO-term analysis among 378 RNAs correlated with L1PA r>0.6 in breast cancer (Raudvere et al., 2019) (Appendix D[E]). Whereas GO-terms were enriched for RNA processing, RNA splicing, RNA metabolism and mitochondrial function among 233 RNAs correlated with L1PA in ovarian cancer (r>0.4) (Appendix D[E]). In this analysis and enriched terms, networks of interacting RNA binding proteins were apparent (Figure 4-2). Remarkably TDP-43, which was previously shown to control retrotransposon activity in multiple studies (Krug et al., 2017; Pereira et al., 2018; Liu et al., 2019), demonstrated in breast cancer the strongest negative correlation with retrotransposon RNAs.

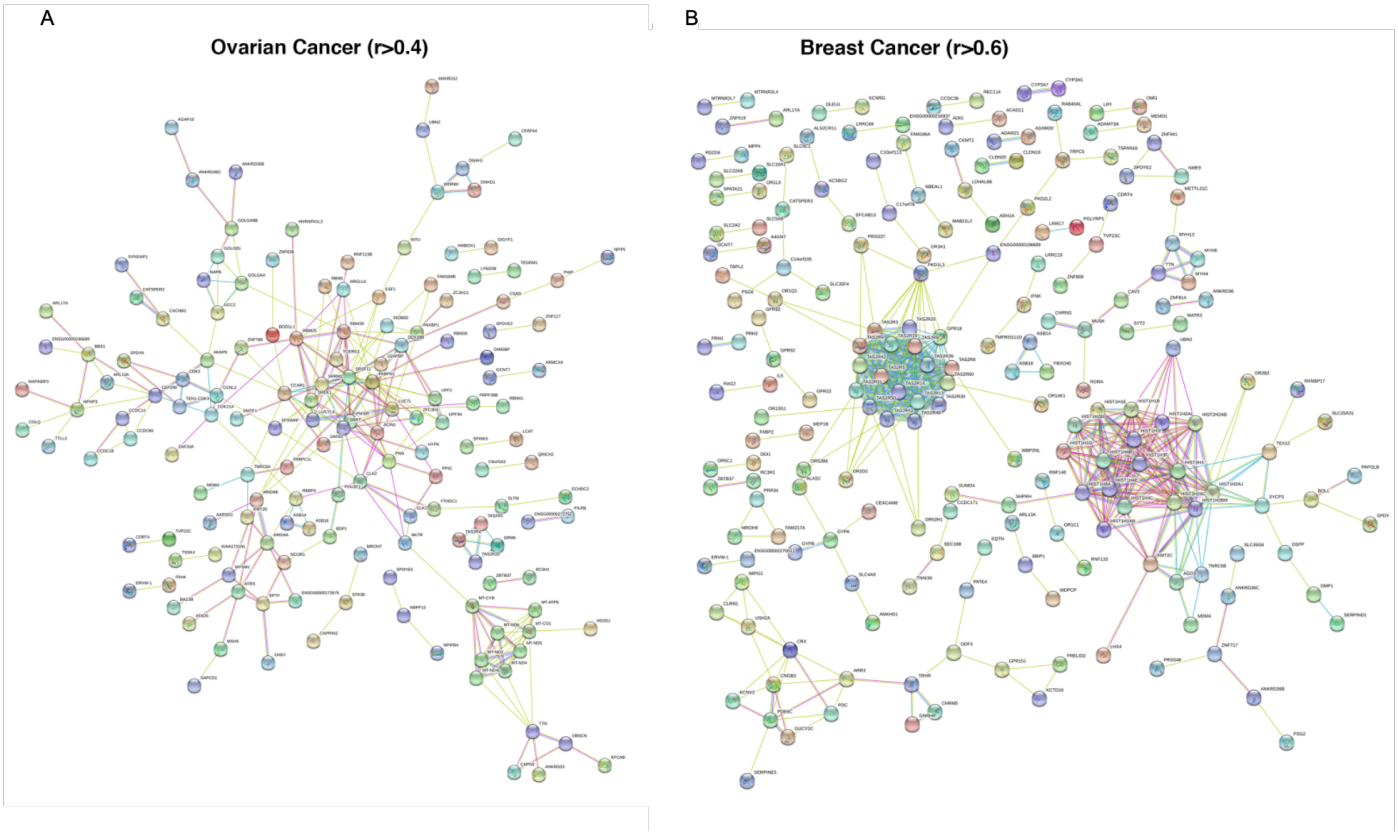


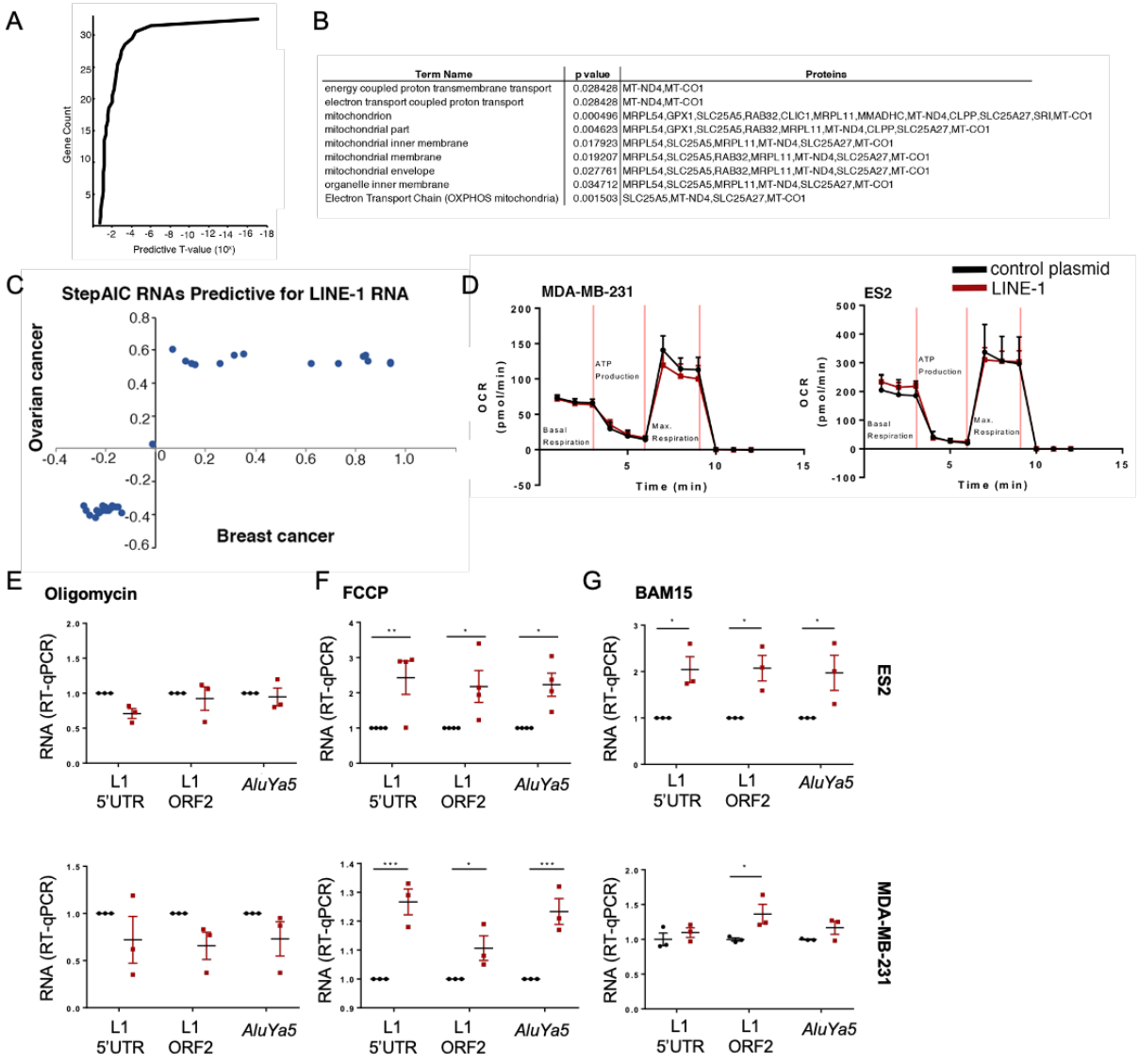
Figure 4-2 Protein interaction network of RNAs correlated with L1 in ovarian and breast cancer
RNAs correlated with L1 r>0.4 (**A.** ovarian cancer) or r>0.6 (**B.** breast cancer). Protein interaction network was generated with STRING excluding data based on co-expression, co-occurrence, gene fusion or neighborhood. Graphs were reproduced with permission from (Alkailani et al., 2021), Copyright Oxford University Press.

To identify RNA features of ovarian cancer with maximum predictive power for expression of retrotransposons, we incorporated StepAIC (Step Akaike Information Criteria) into our analyses. StepAIC is a type of analysis in R software used to simplify our model and select the minimum number of RNAs predictive for L1 expression (Akaike, 1974; Zhang, 2016). In this analysis, thirty-four RNAs provided a notable plateau of predictive power ($t=8x10^{-18}$) for L1PA expression in ovarian cancer (Figure 4-3 A, Appendix D[F]). Decreases in regulators of autophagy and lysosomal degradation (LAMTOR1, ATP6V0E1, RAB1B) were among RNA features associated with elevated L1PA RNA levels in line with a previously observed role of autophagy in retrotransposon

RNA degradation (Guo et al., 2014). Remarkably, factors involved in oxidative phosphorylation, mitochondrial fission and mitochondrial translation constituted 32% of these 34 RNAs (Figure 4-3 B). In addition, GO-terms associated with mitochondria and oxidative stress were the only ones statistically enriched in L1 RNA predictive features for ovarian cancer, and many of them were similarly correlated with retrotransposon expression in breast cancer (Figure 4-3 C). This evidence indicates that oxidative stress or mitochondrial activity is highly predictive of retrotransposon RNA levels in breast and ovarian cancer. We followed a standard protocol measuring cellular oxygen consumption rates (OCR) to evaluate whether retrotransposon expression affects cell culture's mitochondrial activity. No impact was observed of expressing L1 in ES2 and MDA-MB-231 cells on mitochondrial respiration (resting, leak, ATP-linked, maximal respiration rates, or spare capacity) compared to cells transfected with a control plasmid (Figure 4-3 D).

To investigate if mitochondrial activity and oxidative stress affect retrotransposon expression, we treated breast and ovarian cancer cells with drugs to either reduce (oligomycin) or maximize mitochondrial respiration (FCCP (Carbonyl cyanide 4-[trifluoromethoxy]phenylhydrazone)). Although decreasing mitochondrial respiration showed no effect on retrotransposon RNA levels (Figure 4-3 E) in either cell type, maximizing mitochondrial respiration increased L1 and *AluYa5* RNA levels in both cell lines (Figure 4-3 F). To confirm if the observed FCCP action is specific to the mitochondrial activity, BAM15, another protonophore and mitochondrial uncoupler, was similarly employed. Retrotransposon RNA expression was augmented comparably upon BAM15 treatment (Figure 4-3 G). This evidence suggests that augmented mitochondrial

activity to meet the increased metabolic needs in tumors might enhance retrotransposon expression accordingly.



Term Name	p value	Proteins
energy coupled proton transmembrane transport	0.028428	MT-ND4,MT-CO1
electron transport coupled proton transport	0.028428	MT-ND4,MT-CO1
mitochondrion	0.000496	MRPL54,GPX1,SLC25A5,RAB32,CLIC1,MRPL11,MMADHC,MT-ND4,CLPP,SLC25A27,SRI,MT-CO1
mitochondrial part	0.004623	MRPL54,GPX1,SLC25A5,RAB32,MRPL11,MT-ND4,CLPP,SLC25A27,MT-CO1
mitochondrial inner membrane	0.017923	MRPL54,SLC25A5,MRPL11,MT-ND4,SLC25A27,MT-CO1
mitochondrial membrane	0.019207	MRPL54,SLC25A5,RAB32,MRPL11,MT-ND4,SLC25A27,MT-CO1
mitochondrial envelope	0.027761	MRPL54,SLC25A5,RAB32,MRPL11,MT-ND4,SLC25A27,MT-CO1
organelle inner membrane	0.034712	MRPL54,SLC25A5,MRPL11,MT-ND4,SLC25A27,MT-CO1
Electron Transport Chain (OXPHOS mitochondria)	0.001503	SLC25A5,MT-ND4,SLC25A27,MT-CO1

Figure 4-3 Increased mitochondrial activity induces expression of retrotransposon RNA.
A. Plot of gene count vs. predictive power of L1 expression using STEPAIC. **B**. GO-terms significantly associated with 34 STEPAIC genes with predictive power for L1 expression, see Appendix D[F] for the complete set of data **C**. Plot of the correlation values of the 34 STEPAIC genes with L1 in ovarian and breast cancer. **D**. Plot of oxygen consumption rate over time of MDA-MB-231 or ES2 cells transfected with either pcDNA3.1 or a L1 expression plasmid as measured on an Agilent Seahorse. Cells were treated with oligomycin, FCCP and then antimycin and rotenone. Plot represents mean values of at least three independent biological replicates and standard error of the mean. **E-G**. RT-qPCR of L1 and *AluYa5* RNA levels after treatment of cells with 1µg/ml oligomycin (**E**), 0.4µM FCCP (**F**) or 0.1µM BAM15 (**G**) in ES2 cells (upper) and MDA-MB-231 (lower). N=3-4 independent biological replicates. In RT-qPCR horizontal lines represent averages and error bars in all graphs represent standard error of the mean. *p<0.05, **p<0.01, ***p<0.001, two-way ANOVA with Holmes-Sidak correction.

4.2.3 A subset of RNAs correlated with retrotransposons induce their expression

Candidate genes and complexes that could control the expression of broad classes of retrotransposons were selected from the lists above. These include chromatin or DNA-modifying proteins (CHD2, POU5F2) and multiple RNA binding proteins (SETX, PNN, BAT1/DDX39B, RNPC3, SNRNP70, DGCR8 and TNRC6A). All of which were positively correlated with retrotransposon RNA levels, details about each candidate and how they might control retrotransposon expression are included in Table 4-1 and Figure 4-4, respectively. RT-qPCR primers were designed to differentiate retrotransposition-competent L1 (L1HS) from other L1PA elements and *AluYa5* and *AluYb8* from other *Alu* to test the effect of candidates on L1HS and *Alu* family members (Appendix E). Although knockdown of several candidates had no significant impact on the expression of retrotransposon RNA (RNPC3, SETX, DDX39B, TNRC6A, DGCR8, Figure 4-5 A,B), others (PNN, SNRNP70 and CHD2) did significantly decrease expression of L1 and *Alu* RNAs (Figure 4-5 C,D). This evidence suggests that these candidates might impact endogenous retrotransposons expression.

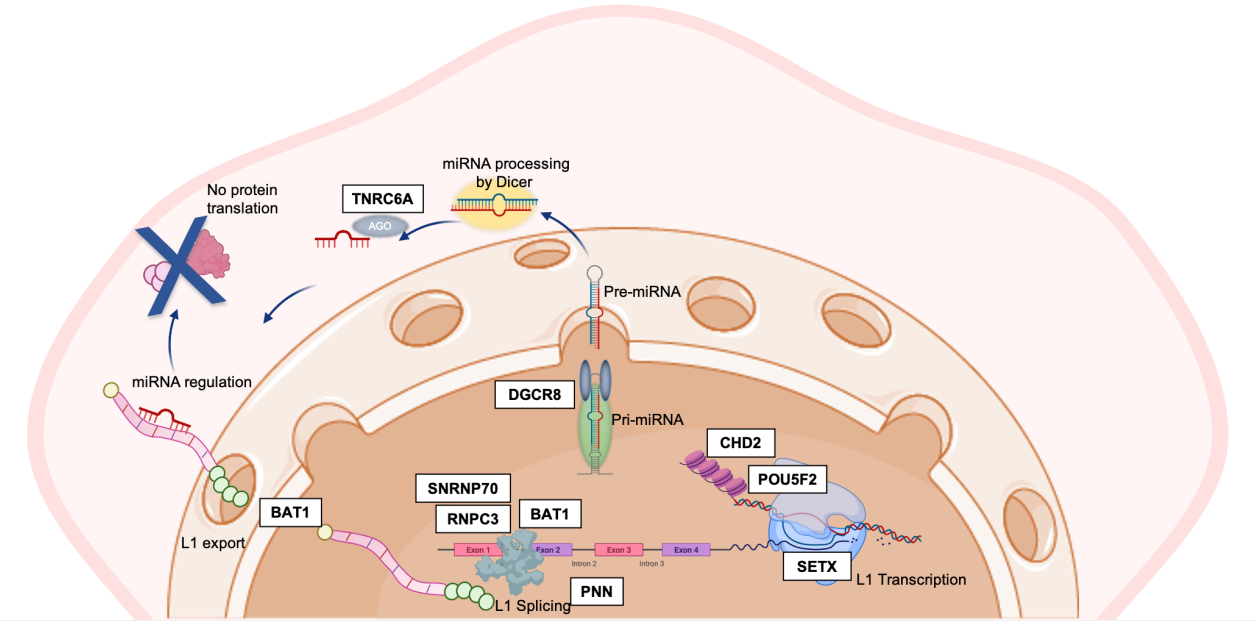


Figure 4-4 Possible regulation mechanisms of candidate genes in the light of L1 life cycle.
Localizations, functions and mechanism of regulation are inferred from related literature cited in Table 4-1 and (Treiber et al., 2019) for miRNA biogenesis process. Illustrations were created with BioRender.

Table 4-1 List of genes positively correlated with retrotransposon RNA and their *in vitro* test results summary.
Summarized effects of candidates' knockdown on retrotransposons are based on data presented in Figure 4-4.

Gene	Function	Subcellular localization (UniProtKB)	References	Effect of knockdown on retrotransposon RNA	Effect of knockdown on insertions (*in vitro*)
PNN	⇒ Tumor suppressor for renal cell carcinoma ⇒ Exon-junction complex associated splicing factor ⇒ Post-transcriptional regulation	Nuclear	(Wang et al., 2002; Alpatov et al., 2004; Murachelli et al., 2012)	Decreased	Decreased
BAT1/ DDX39B/ UAP56	⇒ Splice factor required for the first ATP-dependent step in spliceosome assembly ⇒ Involved in nuclear export of mRNA.	Nuclear & Cytosolic	(Luo et al., 2001; Shen et al., 2008; Linder and Jankowsky, 2011)	Insignificant increase	Decreased
CHD2	⇒ DNA-binding helicase specifically binds to the promoter of target genes, leading to chromatin remodeling by promoting deposition of histone H3.3	Nuclear	(Carvill et al., 2013; Luijsterburg et al., 2016)	Decreased	Increased
RNPC3	⇒ A component of the minor U12-type spliceosome. ⇒ Binds to the 3'-stem-loop of m^7G-capped U12 snRNA	Nuclear & Cytosolic	(Benecke et al., 2005)	Insignificant increase	Decreased
SNRNP70	⇒ mRNA major splicing ⇒ Component of the U1 snRNP	Nuclear & Cytosolic	(So et al., 2016)	Decreased	Decreased

SETX	⇒ Probable RNA/DNA helicase involved in diverse aspects of RNA metabolism and genomic integrity	Nuclear & Cytosolic	(Hatchi et al., 2015)	No significant change	Insignificant decrease
TNRC6A	⇒ miRNA- guided cleavage ⇒ Translational repression by interaction with members of Argonaute family	Cytosolic	(Jakymiw et al., 2005; Liu et al., 2005)	Insignificant increase	Insignificant increase
DGCR8	⇒ Component of microprocessor complex ⇒ miRNA biogenesis ⇒ Bind retrotransposon RNA	Nuclear	(Heras et al., 2013)	Insignificant increase	Insignificant increase
POUF5F2	⇒ DNA binding ⇒ Regulatory function in meiosis	Nuclear	(Nieto et al., 2007)	Not available	Insignificant decrease

We sought to identify the effect of the selected candidates on retrotransposon activity by following a standard *Alu* retrotransposition assay protocol. The used neomycin resistance reporter construct helps avoid reporter-dependent artifacts that may arise when L1 reporters are used in conjunction with siRNA (Cook and Tabor, 2016). The co-transfection may affect the exogenous promoter activity and lead to inconsistent findings with different reporters (Cook and Tabor, 2016). The number of selected neomycin-resistant colonies defines the retrotransposition rate.

Five out of nine candidates affected retrotransposition levels significantly. While the knockdown of PNN, DDX39B, RNPC3 and SNRNP70 decreased *Alu* retrotransposition, CHD2 augmented it (Figure 4-5 E). There is a possibility that these candidate genes affect the transfected plasmid or the splicing events required by this retrotransposition assay but not by endogenous retrotransposons. Also, the observed correlations with retrotransposon RNA include all retrotransposons in the genome, whether embedded within other genes or independent in the genome. Cumulatively, as the knockdown of PNN, SNRNP70, and CHD2 decreased levels of endogenous retrotransposon RNAs and impacted retrotransposition rates significantly, they are likely to direct regulators of endogenous

retrotransposons. Even with taking the previously mentioned caveats into account, our analysis of mRNAs correlated with retrotransposon RNAs can serve as a strategy to identify candidate factors affecting retrotransposons in tumors.

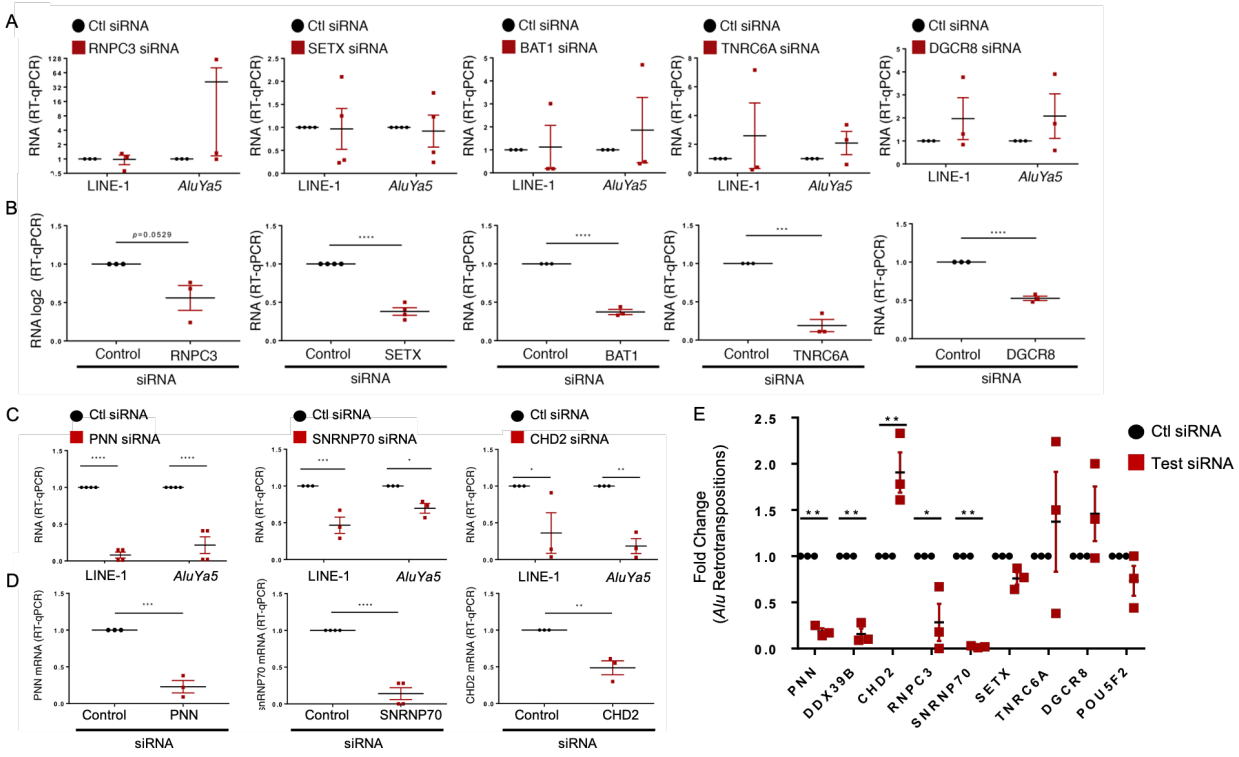


Figure 4-5 A subset of candidate RNAs correlated with retrotransposons induce their expression
A,C. RT-qPCR of L1 (ORF2) and *AluYa5* retrotransposon RNA two days after transfection with indicated siRNA. N=3-4 independent biological replicates. **B,D.** RT-qPCR of the target of the siRNA performed in the samples used in (**A,C**) to validate target knockdown. N=3-4 independent biological replicates. **E.** Alu retrotransposition assay fold-change in colony numbers after transfecting cells with control siRNA or siRNA targeting the indicated test gene. Horizontal lines represent averages and error bars indicate standard error of the mean. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001 two-way ANOVA with Holmes-Sidak correction (**A,C,E**) t-test (**B,D**). Graphs were reproduced with permission from (Alkailani et al., 2021), Copyright Oxford University Press.

4.3 Discussion

Unlike protein-coding genes, retrotransposons are substantially less conserved among different species. This may indicate their non-essentiality to human life; however, the biological activity of retrotransposons is evident in the literature. Therefore, with this low evolutionary conservation, it is challenging to identify retrotransposons whose activity results in biological consequences (functional) among all active elements in humans

(Göke and Ng, 2016). Overcoming these challenges, we pursued various strategies in this study, including using Bayesian correlations, designing specific primers to detect the most active retrotransposon elements, using interactome and GO-terms tools to guide our search for biological effects of retrotransposons and factors regulating them. Unsurprisingly, the observed correlations between active retrotransposons classes were the strongest amongst all other RNAs. This observation could be related but not limited to several reasons. Firstly, the possible shared origin of RT sequences in both LTR and non-LTR (Xiong and Eickbush, 1990). How their shared origin would affect their activity is not entirely clear, they may activate similar pathways. The utilization of L1 machinery by *Alu* and its RNA stabilization by ORF1p is an established possibility (Deininger, 2011). The most likely explanation is that they share a similar network of regulatory mechanisms to repress them; as mentioned in the first chapter, if it is relieved, all of them are activated simultaneously. For example, DNMTi treatment was associated with activating L1, *Alu*, and HERV elements (Roulois et al., 2015; Saito et al., 2016; Kong et al., 2019). The high enrichment observed in binding sites of several transcription factors in breast and ovarian cancers also suggests that a gene expression program controlled by these factors might be involved in the regulation of retrotransposon expression.

DGCR8 acts as a part of the Microprocessor complex that cleaves primary miRNAs (pri-miRNAs) (Heras et al., 2013). These pri-miRNAs are processed by DICER in the cytoplasm to mature miRNAs loaded into Argonaute (AGO)-containing RISCs (Heras et al., 2013). Although our data suggested a positive correlation between DGCR8 and retrotransposon RNA expression, its depletion caused an insignificant increase in the levels of RNA and genomic insertions. This observation could be related to the DGCR8

role previously shown to restrict the activity of L1 and *Alu* in cell culture assays (Heras et al., 2013).

The enrichment of RNAs involved in mitochondrial function among those most correlated with retrotransposons is an exciting observation which calls for further investigation. Our observation that some mitochondrial RNAs were highly correlated while others were negatively correlated with retrotransposons might suggest that cells adapt cellular energy resources to meet the needs of a tumor.

The link between retrotransposon expression and mitochondria is also evident in the literature. In tumor adjacent tissues, mitochondrial genes and immune genes were negatively correlated with intronic TEs (embedded within genes) (Chung et al., 2019). This observation is significant because it was made in control tissues. The observed effect could be related to the regulation of genes (where the TEs are embedded), with no direct impact on TEs. Alternatively, if they impact TEs, mitochondrial and immune RNAs may contribute to TEs restriction mechanisms before the transformation. Our data included a subset of mitochondrial RNAs that were also negatively correlated with retrotransposon RNAs. In addition, our observation that the mitochondrial activity regulates levels of retrotransposon RNA directly was made in cancer, not in normal cells (Figure 4-3 E-G). The mitochondrial regulation of retrotransposon levels may be a part of the homeostatic mechanism that keeps them under certain thresholds.

When the mitochondrial activity is altered due to drug treatment or a disease condition, this homeostasis is disturbed (Fang et al., 2016). This disturbance could lead to changes in the metabolic pool of NAD+, a coenzyme that accepts hydride equivalents to form

reduced NADH in the mitochondrial electron transport chain (Cantó et al., 2015). NAD+ acts as a link between cellular metabolism and transcriptional signaling (Cantó et al., 2015). PARP and sirtuin enzyme families are the primary consumers of NAD+ (Cantó et al., 2015). Therefore, if NAD+ levels are altered due to changes in mitochondrial functions, the activity of these enzymes may be changed, which may impact the retrotransposon expression. This speculation can be the core of an exciting topic for future research.

It is also worth mentioning that mitochondria undergo morphological and metabolic changes during the differentiation of stem cells (Khacho and Slack, 2017). Mitochondria in stem cells start as immature, globular structures with poorly developed cristae, and they develop into the known morphology during differentiation (St. John et al., 2005). They also switch from glycolytic to oxidative phosphorylation metabolic states (Folmes et al., 2011). During stem cell differentiation, retrotransposons are activated with lower epigenetic control over them (Wang et al., 2020). This activation of retrotransposons during stem cell differentiation could also be related to higher levels of mitochondrial activity, which needs further investigation.

In another piece of evidence, a lncRNA called SAMMSON (survival associated mitochondrial melanoma-specific oncogenic non-coding RNA), which has a TSS derived from a solitary LTR1 retrotransposon, was shown to contribute to cancer cell-specific mitochondrial functions in melanoma models (Babaian and Mager, 2016; Leucci et al., 2016). Its upregulation in melanoma could be promoted by its retroviral origin (Göke and Ng, 2016). This evidence suggests retrotransposon control of mitochondria genes, unlike

our observations, which show that mitochondria regulate retrotransposon expression. This difference could be related to tumor type and heterogeneity across different systems. Interestingly, the mitochondrial membrane harbors sensor molecules that trigger antiviral responses called MAVS, mitochondrial antiviral signaling protein (Seth et al., 2005). These act downstream of MDA5 or RIG-I and upstream of NF-κB, triggering the induction of proinflammatory cytokines (Seth et al., 2005). It is possible that the increase we observed in the mitochondrial activity that promoted retrotransposon expression may trigger MAVS induction and immune response consequently.

We used two types of experiments, RNA quantification and *Alu* retrotransposition assays, to test candidate genes. Neither one could distinguish between independent and embedded retrotransposons effect so that the observed effect could be related to one or the other. Despite this limitation, our data provided evidence of regulation that might be useful to investigate further.

The latter assays utilized a reporter system that requires splicing in order to obtain the results. Therefore, the observed effects of some candidates might be due to their role in the splicing of the exogenous reporter system, which would not be relevant to endogenous retrotransposons that are not spliced, except if retrotransposons are embedded in the introns. Then these require splicing for their activation and therefore might be affected by the candidate genes. Three candidates showed significant impact on retrotransposons in both assays (CHD2, PNN and SNRNP70); however, only PNN and SNRNP70 showed a consistent effect. PNN, a member of the exon junction complex required to bind RNAs

during splicing and associate to them following their export to the cytoplasm (Alpatov et al., 2004).

Interestingly, this complex was shown in a *Drosophila* model to restrict TEs by ensuring complete splicing of PIWI transcripts; however, PNN was excluded from having a role in this restriction mechanism (Malone et al., 2014). It is possible that under stress conditions, the exon junction complex cannot control for efficient splicing of PIWI transcripts. At the same time, PNN could support the splicing of retrotransposons in the nucleus by binding their RNA before their export (Figure 4-4).

SNRNP70 was among RNAs shown to associate to ORF1p in ectopically expressed L1 immunoprecipitation assays; however, the consequences of this association were not assessed further (Goodier et al., 2013). Evidence shows that Sm (smith) core protein subunits within the U1 snRNP (U1 small nuclear ribonucleoprotein particle, that is SNRNP70 part of it) are targets for autoreactive B cells and T cells in several rheumatic diseases, including SLE (Kattah et al., 2010). In SLE, the autoimmune response against SNRNP70 is closely linked to TLR and the activation of type-I IFN (Kattah et al., 2010). Recent evidence also indicated the detection of high levels of autoantibodies against ORF1p in SLE patients (Carter et al., 2020). SNRNP70 directly binds ORF1p, and both could associate with retrotransposon RNA in SGs or other cytosolic macromolecular assemblies. In stress conditions, these may trigger the TLR activation of the IFN response. This speculation needs further exploration and may lead to a better understanding of the link between retrotransposons and autoimmune disease.

It would be interesting to follow up on the candidates mentioned above and examine how they regulate retrotransposon expression. More broadly, our strategy to identify mitochondrial regulation of retrotransposons and mRNAs correlated with them might help in surveying candidate factors affecting retrotransposons in tumors or other disease contexts.

4.4 Methods

4.4.1 Transposon expression analysis

The method was described previously with the addition of GO-term enrichment analysis in which Gprofiler was used (Raudvere et al., 2019).

4.4.2 Cells and Reagents

The method was described earlier, with the addition of HeLa (CCL2, ATCC) cells cultured in Dulbecco's Modified Eagle's Medium containing 10% fetal bovine serum and 2 mM L-glutamine. Silencer Select siRNAs (Life Technologies, Table 4-2) were transfected using RNAiMax at a final concentration of 10 nM.

Table 4-2 siRNAs used in chapter 4 identifiers and sequences

siRNA Name	siRNA ID	Sense Sequence	Antisense Sequence
BAT1	s15472	AGUUCAUGCAAGAUCCAAUTT	AUUGGAUCUUGCAUGAACUTG
CHD2	s2978	CGAAUCAGCAGGUUCCAAATT	UUUGGAACCUGCUGAUUCGCT
DGCR8	s29062	GGAUCAUGACAUUCCAUAATT	UUAUGGAAUGUCAUGAUCCAC
PNN	s10757	GGAUACUUCAGGACUAGAATT	UUCUAGUCCUGAAGUAUCCCT
SNRNP70	s13210	CAUGCACUCCGCUUACAAATT	UUUGUAAGCGGAGUGCAUGTC
SETX	s22952	GAGUAGUUGAUACCCGAAATT	UUUCGGGUAUCAACUACUCCA
TNRC6A	s26154	GUCGCAAAAUGGAGAUUGATT	UCAAUCUCCAUUUUGCGACGT
RNPC3	s57832	GGAAUUAUCUAGUGAAGAATT	UUCUUCACUAGAUAAUUCCTC

4.4.3 RT-qPCR

Experiments were performed as previously described in section 3.4.3 with primers listed in Table 4-3 below.

Table 4-3 Oligos used in chapter 4 sequences

Target	Forward Oligos (5'-3')	Reverse Oligos (5'-3')
L1-5'UTR	ACG GAA TCT CGC TGA TTG CTA	AAG CAA GCC TGG GCA ATG
TNRC6A	CTG AGG CTG GCT TAC ACT GG	CGT AGA AAG AAA TCT TAG CAG GG AT
SNRNP70	TCA TCT TGC CAA GTC GCG T	CTT AGC GGG CGA CGG AAT C
RNPC3	AGC TGT TAC GCA CAG TTC CA	CCC ATG GTG GTT CAG TTT GC
DGCR8	ACC ATC AAT GGT CAC CGA GG	TGC ACC CAC AAA GAA AGA GTT TG
SETX-2	TGG TTT TCC ATC AGG GTC CTC	CAG GGT CGG CAG AAG GAT TG
PNN	GGA TGG GCC TCA CGT CAT TC	CGG CCT GTA AAG CAG TCT CA
CHD2	CTG GCT GGG ATT TGG AAC CT	GCA TCG AGT CAC TCA GCC TC
BAT-1	TAC CGG CTC TGC TCG ATG TA	CAC CAA AGG CCT AGC CAT CA

4.4.4 Oxygen Consumption

The Seahorse XF96 Extracellular Flux Analyzer (Seahorse Biosciences; North Billerica, MA, USA) was used to measure OCR in ES2 and MDA-MB-231 cells. Cells were seeded onto 96-well Seahorse plates at a density of 1×10^3 cells per well. Cells were transfected with either L1-neo-TET or pcDNA3.1 plasmid. The sensor cartridge was hydrated with distilled water 24hr before the assay and incubated sealed in CO2-free, 37 °C incubator. Media on cells was changed to previously warmed phenol red-free XF DMEM media supplemented by 2mM sodium pyruvate, 2mM L-glutamine and 25 mM D-glucose and incubated 1hr in CO_2 free, 37 °C incubator before assay. Meanwhile, drug injections were loaded onto the cartridge. During assay running, resting respiration was measured and cells were treated sequentially with the following: oligomycin (10 µM), to measure the nonphosphorylated OCR; FCCP (7.5 µM for ES2 5 µM for MDA-MB-231), to get the

maximal OCR; and antimycin A (10μM) with rotenone (10 μM), to measure the non-mitochondrial OCR. Each measurement was taken over a 5-min interval followed by 5 min mixing. All readings were normalized to the cellular protein content with non-mitochondrial background subtraction.

4.4.5 Alu Retrotransposition Colony Formation Assay

Alu assay was performed as previously (Guo et al., 2014) with modifications. Briefly, 100 × 10^3 HeLa cells were seeded onto 60-mm culture dishes. The following day, Alu-neoTet was co-transfected with the retrotransposition driver plasmid pCEP 5′UTR ORF2-no neo or pcDNA3.1 (empty vector) and 10 nM of siRNA (final concentration) using Lipofectamine 2000 (Invitrogen). One week after transfection, the cells were counted, and 1× 10^5 cells were re-seeded onto 100 mm plates. The next day, 350 μg /ml G418 (Life Technologies) was added and replaced every three days. After ten days of selection, colonies were crystal violet (Sigma-Aldrich) stained and counted.

5 CHAPTER 5: BRCA1 TRANSCRIPTIONAL ACTIVATION OF RETROTRANSPOSONS IN OVARIAN CANCER

5.1 Background

About 65% of women diagnosed with epithelial ovarian cancer die every year worldwide (Lheureux et al., 2019). This high death rate is likely because patients are diagnosed at advanced stages of the disease (Lheureux et al., 2019). High-grade serous tumors bear frequent *TP53* mutations and involve mutations in other genes such as *BRCA1, BRCA2, NF1, RB1* and *CDK12* (Jayson et al., 2014). Furthermore, the HR DNA repair pathway is defective in roughly 50% of these tumors (Jayson et al., 2014). Mutations in *BRCA1* and other double-strand DNA break repair genes are mostly linked to high-grade serous ovarian cancer susceptibility (Lheureux et al., 2019).

BRCA1 export to the cytoplasm has been linked to apoptosis regulation (Fabbro et al., 2004). BARD1 was shown to act as a regulator of BRCA1-dependent apoptosis, and nuclear retention of BRCA1-BARD1 complexes contributes to DNA repair and cellular survival (Fabbro et al., 2004). p53 was shown to play a role in regulating BRCA1 shuttling, and its nuclear retention is increased upon p53 dysfunction (Jiang et al., 2011). BRCA1 subcellular localization is disrupted and becomes predominantly cytoplasmic in cases of cancer (Henderson, 2005). This mislocalization is thought to be due to mutations

in its tandem BRCT repeats (Henderson, 2005). BRCA1 cytosolic mislocalization was suggested to promote breast cancer metastasis (Santivasi et al., 2015).

Being involved in multiple cellular functions and its ability to shuttle between cellular compartments raises the possibility that BRCA1 might regulate retrotransposons directly or indirectly. Recent evidence from two different groups indicates that BRCA1 plays a role in suppressing retrotransposons as part of its HR DNA repair function (Liu et al., 2018; Mita et al., 2020). One of the studies performed a genome-wide CRISPR–Cas9 screen in human chronic myeloid leukemia cells and used a neomycin resistance reporter system (Liu et al., 2018). The other study performed a whole-genome siRNA screen in HeLa cells utilizing a high-throughput microscopy-based retrotransposition assay (Mita et al., 2020). The main focus of both studies was at the retrotransposon insertion level; however, we demonstrated in chapter two that *de novo* L1 insertions minimally impact tumorigenesis or tumor progression in most ovarian cancer cases studied. In contrast, retrotransposon RNA expression predicts better survival rates in ovarian but not breast cancer patients. Tumor grade, size, or invasion differences did not account for this observed difference in survival rates among ovarian cancer patients expressing high L1.

We hypothesized that retrotransposon RNAs are regulating or regulated by factors that contribute to this difference. To better understand the mechanism controlling retrotransposons in ovarian cancer, we examined the correlation with retrotransposon RNAs of genes frequently inactivated in ovarian cancer, including BRCA1.

5.2 Results

5.2.1 Genes Frequently Inactivated in Ovarian Cancer are Controlling Retrotransposon Expression

Genes commonly lost or mutated in ovarian cancer (e.g. NF1, PAX8 and MECOM) (Bell et al., 2011) could control the expression of retrotransposons in a broad subset of patients. Therefore, we sought to evaluate whether their expression correlated with retrotransposon RNA expression. Among these, BRCA1 RNA levels had one of the highest correlations with *Alu* and L1 RNA expression (r = 0.15-0.19, Appendix D[E]), which was less apparent in breast cancer (Appendix D[G]). Up to 23% of ovarian cancer tumors exhibit reduced expression or activity of BRCA1 due to deletion, mutation or gene hypermethylation (Bell et al., 2011). The analyzed ovarian tumors from TCGA, which showed loss of one or more copies of BRCA1 or hypermethylation of its locus, had reduced BRCA1 expression (Figure 5-1 A,B). BRCA1 RNA levels correlated significantly with L1 and *AluY* RNA levels in ovarian tumors (Figure 5-1 C,D). This observation suggests that BRCA1 may control levels of retrotransposon RNA in ovarian tumors.

5.2.2 BRCA1 Supports Retrotransposon RNA Expression

BRCA1 is involved in numerous molecular functions in the cell, including its roles in transcription regulation by binding specific DNA sequences and regulating RNA polymerases I, II and III (Anderson et al., 1998; Veras et al., 2009; Johnston et al., 2016). *In vitro* cultures of spheroids generated from ovarian patients' cancer cells with or without BRCA1 mutations or deletions (as previously described (Pépin et al., 2015)) were used in RT-qPCR to measure RNA encoded by L1HS and *AluY* elements. Spheroid cultures

derived from patients with BRCA1 loss of function demonstrated lower L1 and *Alu* RNA levels than their BRCA1 wild-type counterparts (Figure 5-1 E), supporting our hypothesis that loss of BRCA1 decreases levels of retrotransposon RNA. To dissect the underlying mechanism, we employed different cellular models. Predictably, untransformed mouse ovarian surface epithelial cells with a Cre-lox genetic deletion of a single copy of BRCA1 (Gamwell et al., 2012) demonstrated a 50% reduction in BRCA1 mRNA levels (Figure 5-1 F). *Alu* levels could not be measured in these mouse cells because *Alu* elements are primate-specific. However, L1 RNA levels were significantly decreased in cells that lack a single copy of BRCA1, which mimicked the loss of BRCA1 in patients (Figure 5-1 E,F). Consistently, when BRCA1 was transiently knocked down in ES2 or HeLa cells using either of two independent siRNA sequences, L1 and *AluY* RNA levels were considerably reduced (Figure 5-1 G-I). Northern blots were performed to confirm that the BRCA1 effect is specific to the full-length (6000 kb) competent L1 RNA, utilizing either a single probe targeting the L1HS 5'UTR or multiple probes specific for L1HS. Decreased levels of L1 RNA were also observed by Northern blot when BRCA1 was siRNA depleted (Figure 5-1 J). The data above shows that RNA levels of full-length L1HS, including 5'UTR, ORF1 and ORF2, are decreased when BRCA1 RNA levels are reduced. This evidence suggests that BRCA1 controls levels of retrotransposon RNA that are likely capable of performing retrotransposition.

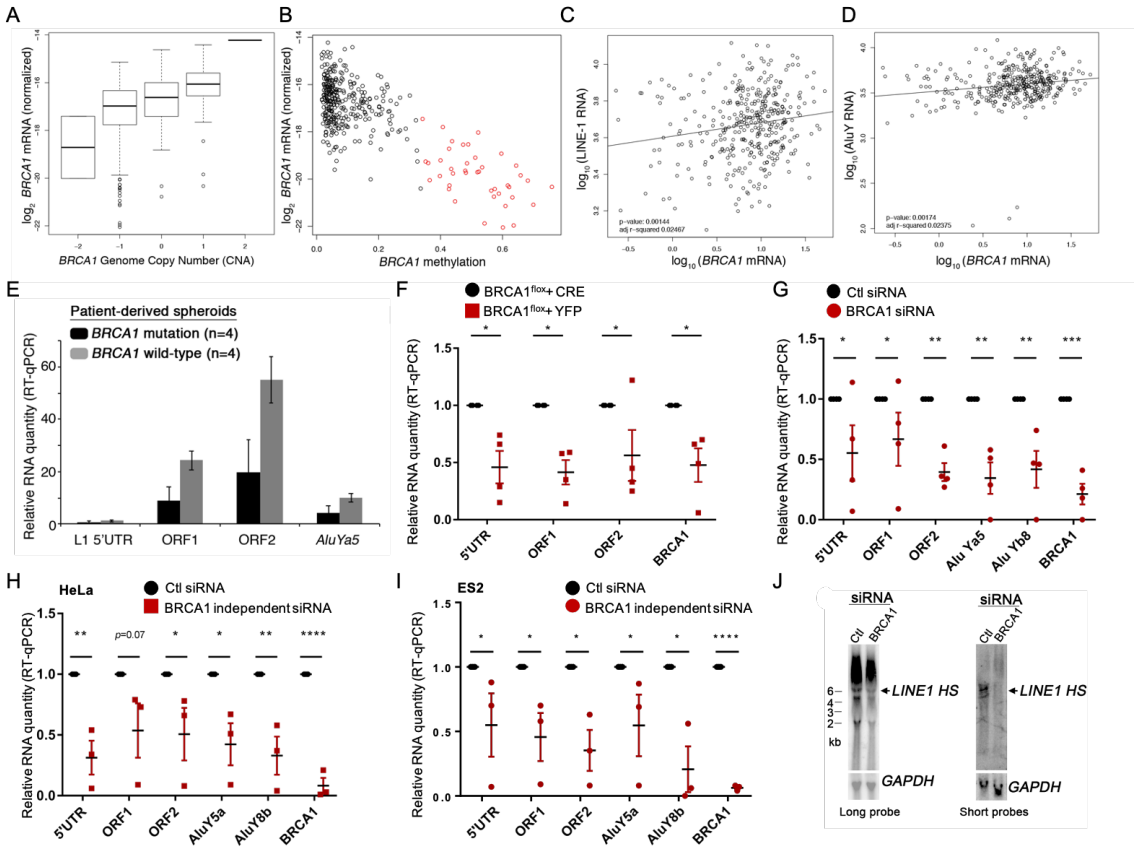


Figure 5-1 BRCA1 controls retrotransposon RNA levels
A. Graph of BRCA1 mRNA levels obtained by RNaseq in tumor samples from TCGA grouped based on the copy number of BRCA1. **B.** Correlation plot of BRCA1 mRNA levels versus methylation scores for BRCA1 in tumor samples from TCGA. Samples with high levels of BRCA1 gene methylation are colored in red. **C,D.** Correlation plot of LINE-1 (**C**) and *AluY* (**D**) RNA levels (proportion of total mapped reads) versus BRCA1 mRNA levels in tumor samples from TCGA. **E**. RT-qPCR of LINE-1 and *AluY* RNA levels in spheroid cultures generated from patient tumors ascites either with or without BRCA1 loss or mutations. Error bars represent standard error of the mean. **F,G**. RT-qPCR of retrotransposon RNA levels in (**F**) mouse primary ovarian epithelial cells wild-type or with loss of one copy of *BRCA1* (Cre-lox) and (**G**) ES2 cells transfected with siRNA targeting BRCA1. **H,I.** RT-qPCR of retrotransposon RNA expression in HeLa (**H**) and ES2 cells (**I**) after transfection with control siRNA or an independent siRNA vs. BRCA1. **J.** Northern blot of LINE-1 in RNA from cells treated with siRNA targeting BRCA1 or control. Blots were probed with either a long probe synthesized with ^{32}CTP that recognizes the 5'UTR of L1HS, or with 12 short ^{32}ATP-labeled probes designed to distinguish L1HS from other L1PA family members. GAPDH was probed as a loading control. Graphs were reproduced with permission from (Alkailani et al., 2021), Copyright Oxford University Press.

5.2.3 BRCA1 Supports Retrotransposon Protein and Genomic Insertions

To evaluate the effect of BRCA1 loss of function on retrotransposon activity, we measured L1 protein levels (ORF1p and ORF2p) in mouse ovarian epithelial cells and ES2 cells by Western blot. Similarly, reduced BRCA1 levels in cells resulted in reduced levels of L1 proteins compared to their controls in either cell model (Figure 5-2 A,B). Furthermore, to

examine if BRCA1 reduction impacts retrotransposon genomic insertion levels, L1 and *Alu* retrotransposition reporter assays were utilized (Moran et al., 1996; Ostertag et al., 2000). In L1 retrotransposition assays, a GFP reporter was used to detect cells with insertion events by flow cytometry, and a version of this reporter with two-point mutations, which ablates retrotransposition, was employed as a negative control. Transient BRCA1 knockdown in ES2 and HEK 293T cells significantly reduced the number of L1 retrotransposition events that occurred (Figure 5-2 C-E). Likewise, significantly fewer *Alu* retrotransposition events were detected when BRCA1 was downregulated with siRNA in ES2 and HeLa cells (Figure 5-2 F,G). Overall, the above evidence suggests that BRCA1 increases L1 and *Alu* retrotransposon RNA, protein and genomic insertions levels (Figure 5-2 H) via a mechanism yet to be investigated.

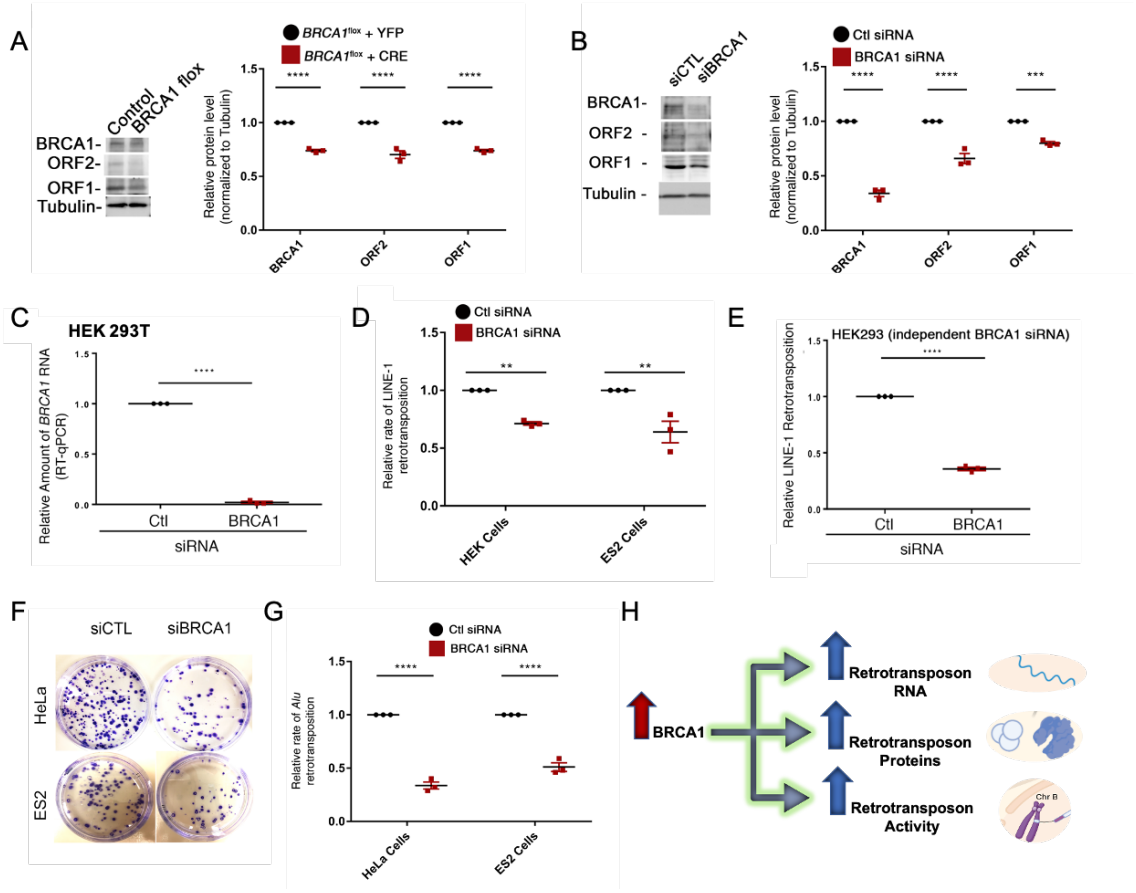


Figure 5-2 BRCA1 controls levels of L1 proteins and retrotransposon genomic insertions
A,B. Western blot of L1 ORF1p and ORF2p proteins in (**A**) mouse primary ovarian epithelial cells wild-type or with loss of one copy of *BRCA1* (Cre-lox) and (**B**) ES2 cells transfected with siRNA targeting BRCA1. Right, quantification of LINE-1 protein levels by western blots N=3 independent biological replicates. **C.** RT-qPCR of BRCA1 mRNA in HEK 293T after transfection with control siRNA or siRNA vs. BRCA1. **D.** Relative rate of L1 retrotranspositions in cells after treatment with control or BRCA1 siRNA. **E.** Relative rates of LINE-1 retrotranspositions in HEK 293T cells after transfection with control siRNA or an independent siRNA vs. BRCA1. **F.** Images of sample plates of HeLa or ES2 cells after selection of neomycin resistant *Alu* retrotransposition colonies from cells treated with siRNA vs *BRCA1* or control siRNA. **G.** Relative rate of *Alu* retrotranspositions in cells after treatment with control or BRCA1 siRNA. **H.** A model to summarize the overall BRCA1 effect on retrotransposon RNA, protein and insertions as represented by data in Figures 5-1 and 5-2, illustrations in **H** were created with BioRender.. In all graphs, horizontal lines represent averages and error bars represent standard error of the mean. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001 two-way ANOVA with Holmes-Sidak correction. Graphs (A-E, G) were reproduced with permission from (Alkailani et al., 2021), Copyright Oxford University Press.

5.2.4 BRCA1 Regulates Retrotransposons by a Mechanism Other Than DNA Damage Repair

L1 reporters were shown to induce DSBs when transiently transfected, and their process of retrotransposition is thought to engage DNA repair machinery (Babushok and Kazazian, 2007). The previously mentioned evidence of BRCA1 suppressing

retrotransposons by repairing the DNA damage during insertion supports the involvement of the HR pathway in the process (Liu et al., 2018; Mita et al., 2020). However, other pieces of evidence on which DNA repair pathway is required for L1 genomic integration are conflicting and, in some cases, are not elucidated enough (Gasior et al., 2006; Babushok and Kazazian, 2007; Suzuki et al., 2009; Coufal et al., 2011). This discrepancy could be due to variations in DNA damage repair pathways between cancer cell models.

As mentioned earlier, BRCA1 is involved in DNA DSB repair mechanisms that contribute to its roles as a tumor suppressor (Wang et al., 2000). To confirm BRCA1 role in controlling levels of retrotransposons by acting through its DNA damage repair function, we measured the effect of depleting essential proteins in either HR or NHEJ pathways as BRCA1 is involved in repairing DNA DSBs mainly through HR and occasionally via NHEJ (Bau et al., 2004; Escribano-Díaz et al., 2013). Transiently silencing Rad50, Rad51, XRCC2, DMC1, MRE11, and XRCC6 by 70-95% with siRNA was insufficient to reduce L1 or *Alu* RNA levels and phenocopy the observed effect of BRCA1 silencing (Figure 5-3 A-C). This observation contradicts previous evidence and indicates that BRCA1 likely controls retrotransposon RNA expression by a mechanism other than DNA repair pathways.

5.2.5 BRCA1 Does Not Stabilize Retrotransposon RNAs by Interacting with ORF1p or Retrotransposons RNA

When BRCA1 was depleted, ORF1p and RNA of both L1 and *Alu* were reduced, suggesting that ORF1p may bind and stabilize L1 and *Alu* RNAs. In an attempt to test this

hypothesis, immunoprecipitation experiments were performed in HeLa cells overexpressing HA-tagged ORF1p. BRCA1 was not detected in HA-ORF1p immunoprecipitates, and HA-ORF1p was not detected in immunoprecipitates of BRCA1 (Figure 5-3 D). This observation suggests that BRCA1 does not bind ORF1p in its regulation of retrotransposon RNA levels.

Evidence has demonstrated that BRCA1 can bind RNA (Ganesan et al., 2002) potentially including L1 RNA (Mita et al., 2020). BRCA1 immunoprecipitation was performed in conditions retaining bound RNA to examine in our system if BRCA1 binds and stabilizes retrotransposon RNA. As a positive control, a known RNA-binding protein (HuR) (López De Silanes et al., 2004) was immunoprecipitated, and RNA was pulled down with it. No retrotransposon RNA was enriched in BRCA1 immunoprecipitates over the IgG control immunoprecipitates (Figure 5-3 E-G), suggesting that BRCA1 does not directly bind retrotransposon RNA to regulate its levels.

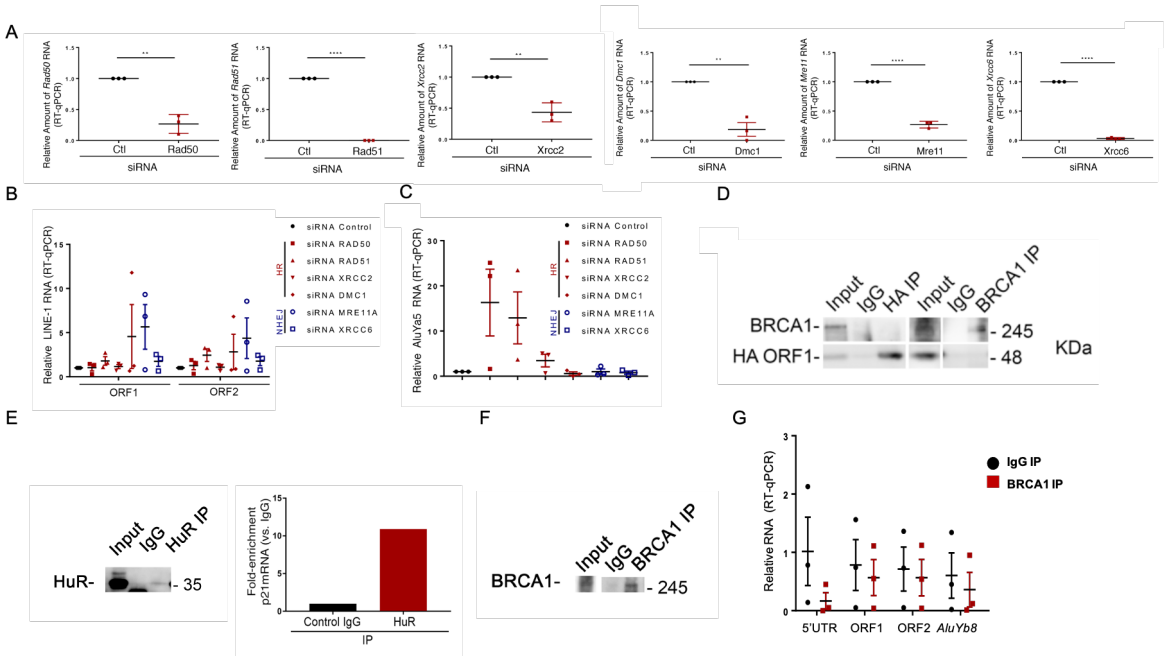


Figure 5-3 BRCA1 affects retrotransposition independent of DNA damage repair, protein interaction or RNA stabilization

A. RT-qPCR of indicated target genes after transfection of cells with siRNA targeting them. N=3 independent experiments. **B,C.** RT-qPCR of LINE-1 (**B**) and *AluYa5* RNA (**C**) two days after transfection of cells with siRNA targeting the indicated genes involved in DNA damage repair. **D**. Western blot of immunoprecipitation of BRCA1 or HA in cells transfected with HA-LINE-1 ORF1. **E.** Western blot of HuR after immunoprecipitation of HuR from HeLa cells (left) and RT-qPCR for HuR-bound mRNA p21 in HuR immunoprecipitated vs. control IgG immunoprecipitated. **F.** Western blot of BRCA1 after immunoprecipitation of BRCA1. **G.** RT-qPCR for the indicated retrotransposon RNAs in BRCA1 IP vs. IgG control IP. In RT-qPCR horizontal lines represent averages and error bars indicate standard error of the mean. *p<0.05, **p<0.01, ***p<0.001, ***p<0.0001, t-test (**A**), two-way ANOVA with Holmes-Sidak correction (**B,C,G**). Graphs were reproduced with permission from (Alkailani et al., 2021), Copyright Oxford University Press.

5.2.6 BRCA1 Supports Transcription of Retrotransposons by Binding to Genomic L1 Copies

We demonstrated above that BRCA1 increases retrotransposon RNA, protein and genomic insertion levels by a mechanism other than DNA damage repair or RNA stabilization. Among other functions of BRCA1, it has been shown to bind chromatin and promote transcription of RNA by RNA polymerase I, II or III (Anderson et al., 1998; Veras et al., 2009; Johnston et al., 2016). Therefore, we sought to examine whether BRCA1 increases retrotransposon RNA levels by acting at the chromatin level and promoting RNA transcription. Public ChIP data (made available from B cells, Embryonic

Stem Cell Line H1 and HepG2) was first analyzed to assess if BRCA1 binds to DNA encoding and surrounding L1. BRCA1- chromatin binding peaks were mapped across the 146 active and intact L1 elements deposited in L1Base (Penzkofer et al., 2017). Intriguingly, peaks of BRCA1 binding were observed within or in proximity to several active L1 elements (Appendix C). Therefore, we performed BRCA1 ChIP in cell models to experimentally confirm this observation. By including the required ChIP controls, chromatin fragments were confirmed to be mostly below 250 bp, and BRCA1 immunoprecipitation was validated by Western blot (Figure 5-4 A,B). Enrichment of a transcriptional pause site in β-Actin mRNA (known to be bound by BRCA1) was detected in BRCA1 ChIPs, whereas no enrichment was detected of an intronic region of β-Actin (not regulated by BRCA1) (Hatchi et al., 2015) (Figure 5-4 C,D). Primers specific to 5', mid and 3' regions of L1 DNA were designed and used to examine if BRCA1 binds to the genomic regions of L1 (Figure 5-4 E). BRCA1 binding to L1 DNA at all tested regions was significantly enriched in BRCA1 ChIPs compared to control IgG counterparts in ES2 and HeLa cell lines (Figure 5-4 F). This data suggests that BRCA1 binds genomic copies of L1 to promote their RNA transcription.

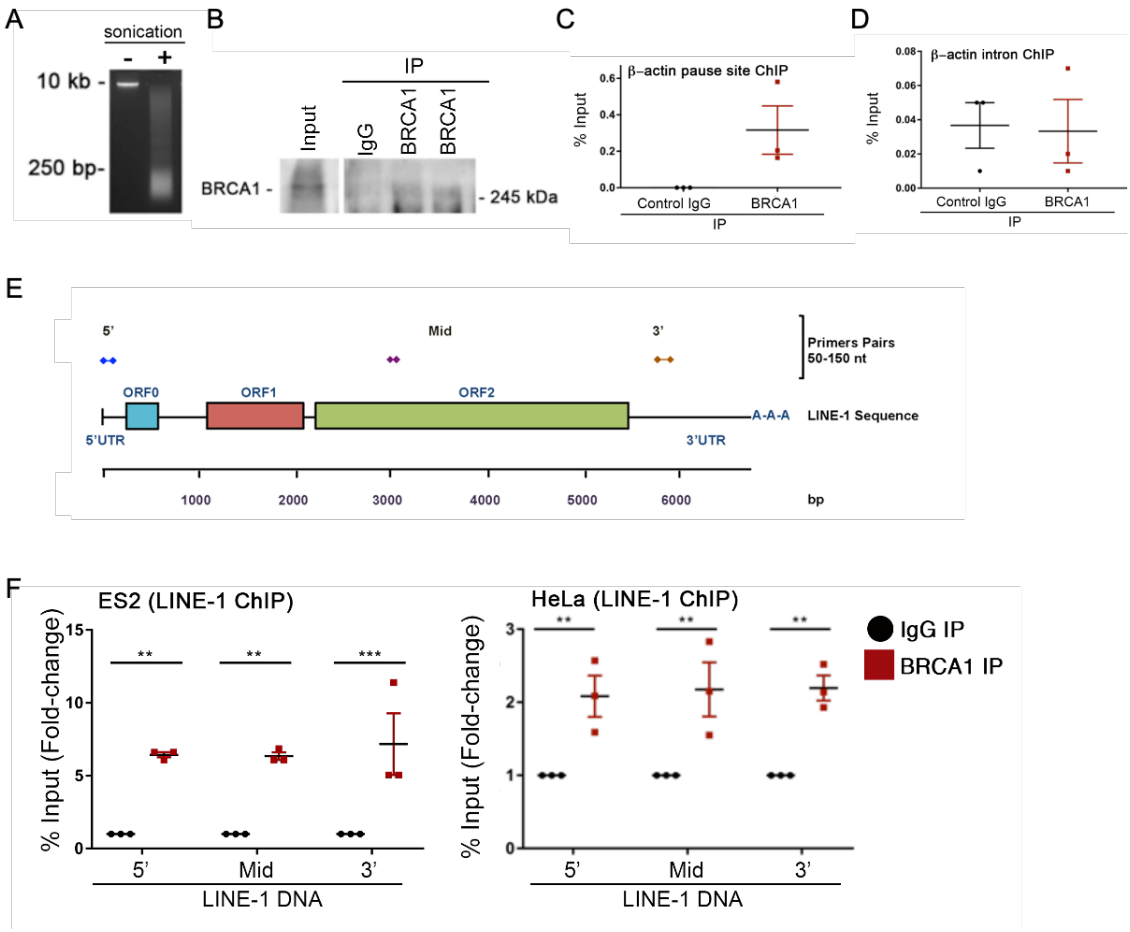


Figure 5-4 BRCA1 associates with DNA copies of L1 in ChIP
A. Chromatin is disrupted by sonication into ~200 bp fragments as expected. **B.** Western blot of BRCA1 immunoprecipitates vs. IgG control immunoprecipitates. **C,D.** Quantitative PCR results of BRCA1 ChIP for genomic sites bound by BRCA1 (β-actin pause site, **C**) or not (β-actin intron, **D**) expressed as percent of input sonicated DNA. **E.** Diagram indicating placement of primers in 5'UTR, mid-region or 3'UTR of LINE-1 for BRCA1 ChIP. **F.** Quantitative PCR for LINE-1 genomic regions normalized to percent input. In all graphs, horizontal lines represent averages and error bars represent standard error of the mean. *p<0.05, **p<0.01, ***p<0.001, two-way ANOVA with Holmes-Sidak correction except (**C,D**) where t-test was used. Graphs were reproduced with permission from (Alkailani et al., 2021), Copyright Oxford University Press.

Nascent RNA was labelled with a pulse of ethynyl-uridine and chased for one hour to confirm the specific effect of BRCA1 on transcription in a biochemical assay. RNA was biotinylated (clicked to ethynyl-uridine), captured on streptavidin-coated beads and quantified. Data were normalized to GAPDH mRNA, which has a half-life exceeding twelve hours across experimental systems (Tani et al., 2012) (Figure 5-5 A). β-actin mRNA, whose transcription is promoted by BRCA1 (Hatchi et al., 2015) accumulated

less in chase samples when BRCA1 was depleted from cells (Figure 5-5 B), validating the utilized pulse-chase assay and the effect of BRCA1 on RNA transcription. BRCA1 knockdown significantly impeded the accumulation of nascent L1 and *Alu* RNAs (Figure 5-5 C). Overall, the data imply that BRCA1 binds genomic regions of retrotransposons and increases the transcription of L1 and *Alu* RNA, thereby promoting their retrotransposition (Figure 5-4, Figure 5-5 A-C).

5.2.7 BRCA1 Regulates Transcription of Retrotransposons Independent of its R-loops Resolution Function

BRCA1 is also involved in resolution of R-loops; these are DNA-RNA hybrids that are often formed near promoters and slow transcription at pause sites (Hatchi et al., 2015). DNA-RNA Immunoprecipitation (DRIP) experiments were performed to examine whether BRCA1 promotes L1 and Alu RNA transcription by helping resolve R-loops. A cocktail of restriction enzymes was used to fragment DNA then R-loops were pulled down using the S9.6 monoclonal antibody. R-loops were enriched in RPL13A (a site known to contain abundant R-loops), and enrichment was eliminated when RNase H was used, which degrades specifically these hybrids, and validates the immunoprecipitation of R-loops (Sanz and Chédin, 2019) (Figure 5-5 D). EGR1, a site known to contain fewer R-loops, was poorly enriched in RNA-DNA hybrids (Sanz and Chédin, 2019) (Figure 5-5 D). Upon BRCA1 depletion, R-loops accumulated at RPL13A, confirming its role in resolving them (Figure 5-5 E). RNA-DNA hybrids were also observed accumulating at L1 genomic regions and abolished by RNase H treatment (Figure 5-5 F). Although BRCA1 siRNA downregulation augmented RNA-DNA hybrids at RPL13A, R-loops at

L1 genomic regions were significantly reduced (Figure 5-5 E,F), suggesting that BRCA1 regulates retrotransposon transcription-independent from its known roles in resolving RNA-DNA hybrids.

Our DRIP assay does not differentiate between R-loops occurring during L1 transcription from those happening during L1 reverse transcription and insertion. If BRCA1 was to promote L1 transcription by resolving R-loops, we should observe higher levels of R-loops pulled down in our assay when BRCA1 is depleted. However, we observed that less R-loops accumulate at L1 genomic regions when BRCA1 was depleted. This suggests that BRCA1 depletion reduces the amount of L1 transcription, and leads to fewer RNA-DNA hybrids being formed during L1 reverse-transcription. Altogether, BRCA1 promotes transcription of retrotransposon RNA (Figure 5-5 G), which leads to more ORF2p derived reverse-transcription, and more R-loops generated.

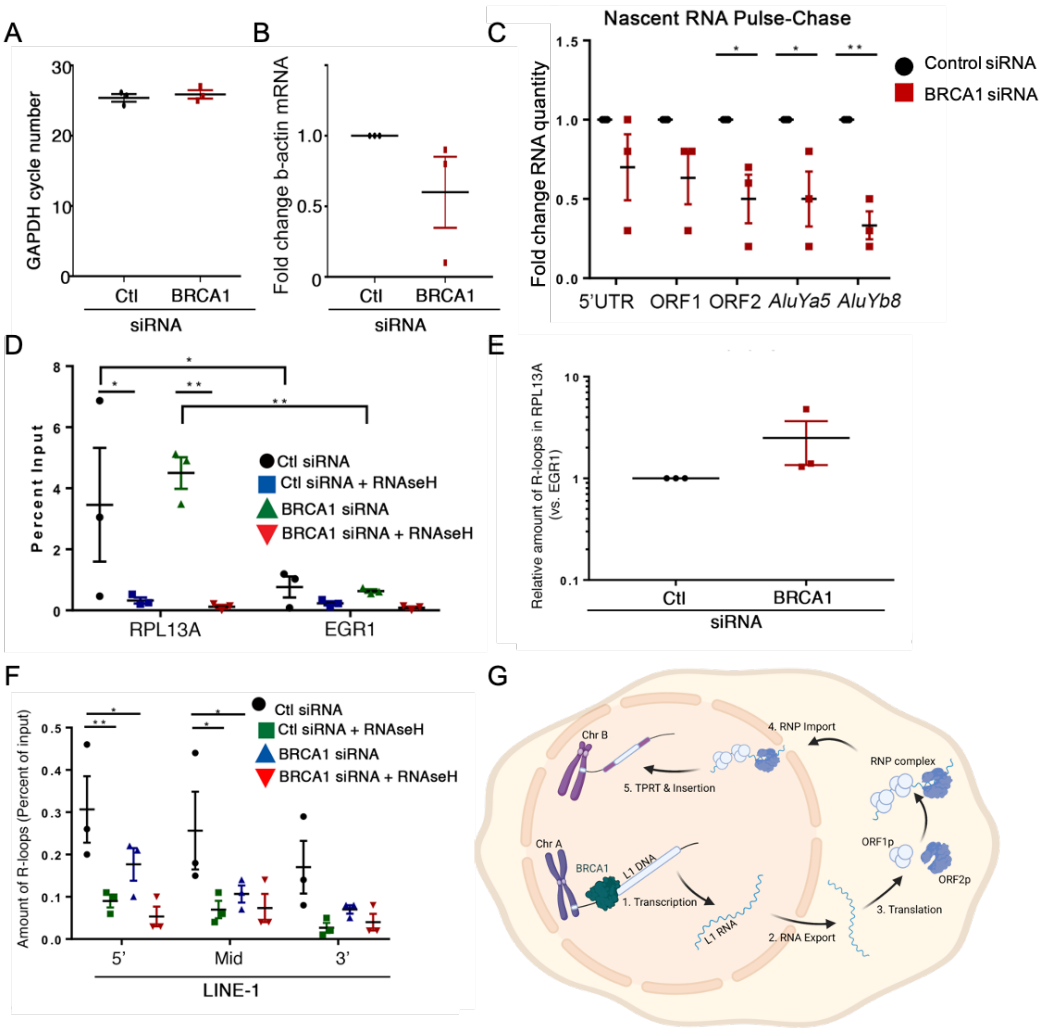


Figure 5-5 BRCA1 Promotes Transcription of Retrotransposon RNA
A-C Nascent RNA pulse-chase analysis in ES2 cells. RT-qPCR of GAPDH crossing threshold cycle number (**A**) and fold-change in β-actin mRNA levels (vs. GAPDH) (**B**), and LINE-1 and *AluY* RNAs among purified nascent RNAs 1 h after pulse of cells with ethynyl-uridine (**C**). **D-F** RT-qPCR of R-loops purified with S9.6 antibody from cells transfected with control or BRCA1 siRNA. **D-F**. RT-qPCR of S9.6 antibody immunoprecipitates (R-loops) in cells transfected with siRNA targeting BRCA1 or control, and S9.6 immunoprecipitates treated or not with RNAseH to degrade RNA-DNA hybrids. RPL13A region accumulate R-loops while EGR1 has few R-loops (**D**). R-loop enrichment is expressed as quantitative PCR values for primers at RPL13A vs. values for a site with few R-loops (EGR1) (**E**). Quantitative PCR of R-loops enrichment at L1 genomic regions in transfected cells (**F**). **G.** A model to represent BRCA1 role in promoting L1 transcription by binding to its genomic DNA in the light of L1 life cycle, model was created with BioRender. Graphs (A-F) were reproduced with permission from (Alkailani et al., 2021), Copyright Oxford University Press.

5.3 Discussion

Investigating regulators of retrotransposon expression in ovarian cancer revealed an interesting mechanism related to BRCA1. Data demonstrated above shows that BRCA1 suppression triggered significant decreases in retrotransposon RNA, protein and genomic

insertion levels. We identified a novel role of BRCA1 in promoting L1 transcription by binding its genomic DNA. In contrast, BRCA1 was demonstrated to suppress L1 retrotransposition by affecting DNA repair events involved in L1 integration into the genome (Liu et al., 2018; Mita et al., 2020). Our study approach is different from these studies' approaches in various ways that might account for the contrasting findings. Doxycycline inducible L1 overexpression plasmids were used throughout the other studies. Therefore, their focus was not to investigate L1 regulation at the transcriptional level. Also, doxycycline treatment was shown to affect mitochondrial translation that shares some sensitivities with bacteria, which yields widespread consequences on cellular metabolism and transcriptional regulation (Houtkooper et al., 2013; Moullan et al., 2015). Our evidence in chapter four also demonstrates that mitochondrial activity regulates retrotransposon expression. Therefore, doxycycline treatment might impact retrotransposon activity due to alterations in the cellular metabolic state. On the other hand, our focus was to study the regulation of retrotransposon RNA produced endogenously in patient tumors, tumor-derived spheroids, primary ovarian epithelial cells and multiple cancer cell lines.

In addition, the other studies used CRISPR to eliminate BRCA1 production. In contrast, our approach included various physiologically relevant tools ranging from endogenous variation in *BRCA1* expression and mutated *BRCA1* in patient spheroids to single-copy deletion of *BRCA1* and siRNA-mediated knockdown in cells. The complete elimination of BRCA1 function may be required to observe the effects of BRCA1 on DNA repair

mechanisms involved in retrotransposon insertions in the genome. i.e., a single copy of BRCA1 is perhaps enough to perform DNA damage repair at retrotransposition sites.

Evidence indicates that L1 is restricted in mouse stem cells by the RING domain E3 ubiquitin ligase (UBR2)- mediated ubiquitination of ORF1p (MacLennan et al., 2017). Also, ORF2p possesses a ubiquitination prone motif at its K842 residue as in the iPTMnet database (Huang et al., 2018). There is a possibility that BRCA1 as an E3 ubiquitin ligase can target L1 proteins for ubiquitination. However, the observed effect of BRCA1 promoting retrotransposon RNA expression, protein production and genomic insertions suggested that L1 proteins were not targeted for degradation. Therefore, we did not test the possibility of BRCA1-mediated ORF1p and ORF2p ubiquitination in this study.

BRCA1 regulation of retrotransposon transcription is not entirely surprising; previous evidence showed BRCA1 regulates transcription (Gardini et al., 2014; Hatchi et al., 2015). By its regulation of RNA polymerases I, II and III (Anderson et al., 1998; Veras et al., 2009; Johnston et al., 2016), BRCA1 could regulate the transcription of independent L1 and *Alu* elements, whose transcription mechanisms are derived by polymerase II and III respectively (Deininger, 2011; Kazazian Jr. and Moran, 2017). BRCA1 could also indirectly regulate the transcription of *Alu* elements embedded in other mRNA transcripts (within introns or 3'UTRs) by modulating their RNA polymerase II (Chiba and Parvin, 2002; Deininger, 2011). Therefore, measuring BRCA1 regulation of *Alu* transcription is challenging; in part, the effect could be related to the mRNA function (where *Alu* is embedded) not directly related to *Alu,* and the transcription of independent elements

occurs at low rates in cells (Deininger, 2011). BRCA1 regulation of L1 transcription in this study suggests that the effect observed is directly related to L1 elements independent of other mRNAs. Overall, the described regulation mechanism may only affect a subset of retrotransposons due to the diversity of their sequences and the sequences surrounding them (Rodriguez-Martin et al., 2020).

Our presented evidence suggests that BRCA1 binds the body of L1 DNA independent of its endogenous promoter TSS (Figure 5-4 F). This evidence is consistent with BRCA1 role in regulating the elongating RNA polymerase II by binding to COBRA1 (Kwak and Lis, 2013). BRCA1 possibly associates to L1 via its N-terminal region as a part of the holo-pol complex that involves BARD1. This regulation process is independent of BRCA1 role in resolving transcription-associated R-loops (Figure 5-5 D-F). Details about what region of BRCA1 is essential for L1 transcription are still unknowns. Introducing Mutant BRCA1 domains into ChIP experiments will help resolve this but was not included in our study.

Reduced expression or activity of BRCA1 due to deletion, mutation or gene hypermethylation affects about 23% of ovarian cancer patients (Bell et al., 2011). One of our study limitations was not to include an analysis of retrotransposon expression in patients based on their *BRCA1* gene state. Multiple common polymorphisms were shown to occur in the *BRCA1* gene, which generate amino acid substitutions (Dunning et al., 1997). A recent study used saturation genome editing to assay for most single-nucleotide variants that might occur in *BRCA1* exons to identify about 300 variants (Findlay et al.,

2018). Although these variants of BRCA1 are of uncertain significance and do not impact the overall risk of developing ovarian cancer (Dunning et al., 1997; Findlay et al., 2018), they may impact the protein function regulating retrotransposons. BRCA1 copy number and methylation rates analyses were performed on patient datasets, and the observed gene expression pattern was as expected, thus analyzing retrotransposon expression for each of BRCA1 known variants could be interesting for future research.

Another limitation was that the cell models used to study the mechanism of BRCA1 regulation of retrotransposons were not entirely of high-grade serous ovarian cancer origin. ES2 cell line is of clear cell carcinoma origin; however, its molecular and phenotypic characteristics are like high-grade serous ovarian cancer (Tudrej et al., 2018). KURAMOCHI, OVCAR-4, and OVPA8 cell lines could be better to resemble the high-grade serous ovarian cancer characteristics (Domcke et al., 2013; Tudrej et al., 2018). In terms of BRCA1 status, ES2 cell line does not possess *BRCA1* mutation, but includes several polymorphic variants of it (2196G > A, 2201C > T, 2430T > C, 2731C > T, 3232A > G, 3667A > G, 4427T > C, 4956A > G) (Stordal et al., 2013). These occur mainly in exon 11, one variant in exon 13, and one in exon 16 with unconfirmed protein effects (Stordal et al., 2013). Since BRCA1 protein function is not certainly affected, ES2 could be used as a cell model in this study. Cell models such as CAOV3 and SKOV3, although they are not included in the top-ranking high-grade serous ovarian cancer cell lines, can serve as a better model than ES2 based on BRCA1 status (Domcke et al., 2013; Stordal et al., 2013). It is worth mentioning that since we observed consistent effects of BRCA1 on retrotransposons in multiple cell lines (ES2, HeLa, and HEK 293T), the mechanism could

be independent of high-grade serous ovarian cancer and is applicable in different cell models.

In general, multiple lines of evidence suggest retrotransposons are involved in the course of tumorigenesis. Our results from high-grade serous ovarian cancer patients suggest that high levels of L1 RNA are associated with improved survival. This evidence is likely not related to a single mechanism *per se* but to a network of interconnected factors regulating and regulated by retrotransposons.

5.4 Methods

5.4.1 RT-qPCR

RT-qPCR was performed as previously described in section 3.4.3. cDNA was prepared in Figure 5-1 using Oligo(dT) and dNTPs with M-Mulv RT (NEB#M0253), and in Figures 5-3 and 5-5 using MiScript II Reverse Transcriptase system (Qiagen). Some of the primers used were described in earlier chapters and the ones listed in Table 5-1 below.

Table 5-1: RT-qPCR primers specific to human or mouse used in Chapter 5.

Human Primer Sets		
Target	**Forward Oligos (5'-3')**	**Reverse Oligos (5'-3')**
L1-ORF1	TCA AAG GAA AGC CCA TCA GAC TA	TGG CCC CCA CTC TCT TCT
Alu-Yb8	ATC CTG GCT AAC AAG GTG AAA CCC	CGG ACT GCT GGA CTG CAG TG-3
BRCA1	TTC ACC AAC ATG CCC ACA GA	CTG CCC AAT TGC ATG GAA GC
MRE11	GAG GAA GAG CAG ACA CTG GT	GAC ATT TCG GGA AGG CTG CT
DMC1	CTC ATA CCC TCT GTG GTG AAC A	ACG GCC ACT GAA ATC CAC TC
RAD50	TGC TTG TTG AAC AGG GTC GT	TCA CTG AAT GGT CCA CGC TC
XRCC6	TGA CAT GAG CAT CCA GTG TAT C	TCC AGC TCC TGT AAG ACG TAA
XRCC2	GCG ATG TGT AGT GCC TTC CA	TCA AGA ATA TCA CCA TGC ACA GG
RAD51	GCT GGG AAC TGC AAC TCA TC	ATT GCC ATT ACT CGG TCC GC
Mouse Primer Sets		
Target	**Forward Oligos (5'-3')**	**Reverse Oligos (5'-3')**
5'UTR	TAA GAG AGC TTG CCA GCA GAG A	GCA GAC CTG GGA GAC AGA TTC T
ORF1	ACT CAA AGC GAG GCA ACA CTA GA	GTT CCA GAT TTC TTT CCT AGG GTTTC
ORF2	CTG GCG AGG ATG TGG AGA A	CCT GCA ATC CCA ACA AT
BRCA1	ACC CCA AGG TCA AGT TGC TT	CGG TCA TTA CTT CTT TGG GGG T
GAPDH	ATG GCC TTC CGT GTT CCT ACC	GTA GCC CAA GAT GCC CTT CAG
β-actin	ACT GCC GCA TCC TCT TCC TC	GGA TGC CAC AGG ATT CCA TAC C

5.4.2 Cells and Reagents

HEK 293T cells (CRL-3216, ATCC) were cultured in Dulbecco's Modified Eagle's Medium containing 10% fetal bovine serum and 2 mM L-glutamine. MOSE cells (from Dr. Barbara Vanderhyden) were cultured in "MOSE medium" as described in (Gamwell et al., 2012). Lipofectamine 2000 (Invitrogen) was used to transfect previously described plasmids or the following plasmids in cells: pEGFP-N1 (a gift from S. Pfeffer, Strasbourg), 99 RPS-GFP PUR, 99 RPS-GFP JM111 PUR, and HA-ORF1 (gifts from J.

Goodier). Sequences of Silencer Select siRNAs (Life Technologies) used are included in Table 5-2 below.

Table 5-2 siRNA used in Chapter 5 identifiers and sequences.

siRNA Name	siRNA ID	Sense Sequence	Antisense Sequence
BRCA1	s457	GGGAUACCAUGCAACAACAUAATT	UUAUGUUGCAUGGUAUCCCTC
BRCA1	s458	CAGCUACCCUUCCAUCAUATT	UAUGAUGGAAGGGUAGCUGTT
RAD50	s793	GCAGUUGUCCAGUUACGAATT	UUCGUAACUGGACAACUGCTC
RAD51	s531930	AGAAGGAGCUAAUAAAUAUTT	AUAUUUAUUAGCUCCUUCUTT
DMC1	s21976	GGACAUUGCUGAUCGCUUUTT	AAAGCGAUCAGCAAUGUCCCT
MRE11	s8960	CCCGAAAUGUCACUACUAATT	UUAGUAGUGACAUUUCGGGAA
XRCC2	s531568	CACUUACUUCUUACACUUUTT	AAAGUGUAAGAAGUAAGUGGG
XRCC6	s52594	GAUCCAGGUUUGAUGCUCATT	UGAGCAUCAAACCUGGAUCAT

5.4.3 Western Blotting

After lysis (RIPA buffer), proteins were resolved on 8% (w/v) acrylamide gels, transferred to PDVF membrane (IPVH00010, EMD Millipore), blocked with 5% milk in TBST (1 hr) and probed with the primary antibody in TBST overnight at 4 °C. The primary antibodies listed in Table 5-3 below were detected with anti-IgG-HRP and HRP substrate (WBLUR0100A, EMD Millipore) using the ImageQuant LAS 4000 system (GE Healthcare). Quantification of blots was performed using the total intensity of an area encompassing the maximum band size using the Image Studio Lite Ver 5.2 software (LI-COR Biosciences) and normalized to Tubulin subtracting background.

Table 5-3 Antibodies used in Chapter 5 identifiers and dilutions.

Antibody	Cat #, Vendor	Dilution
BRCA1	D-9, sc6954, Santa Cruz	1:200
ORF1	Clone 4H1 MABC1152, Millipore	1:1000
ORF2	ETL001, Kerafast	1:250
Tubulin	T9026, Sigma	1:1000
HA	05-904, Milipore	1:1000
HuR	3A2 sc5261, Santa Cruz	1:1000

5.4.4 Cancer Spheroid In Vitro Cultures

A panel of primary ovarian cancer cell lines, derived from ascites of patients with high-grade serous histology (IRB-approved protocol #2007P001918/MGH), with confirmed BRCA1 mutations (N=4) or wildtype (N=4) status by clinical SNaPshot™ genotyping, was used. Briefly, low-passage primary cancer cells were grown in RPMI1640 (Gibco) with B27 supplement (Life Technologies), penicillin/streptomycin (Life Technologies), without any serum additives as previously described (Pépin et al., 2015).

5.4.5 L1 Retrotransposition Assay

The L1 assay was performed as previously (Guo et al., 2014) with modifications. Briefly, 1- 2.5 × 10^5 HEK 293T or ES2 cells per well were seeded in six-well plates. The next day, 1µg of 99 RPS-GFP PUR (L1-WT) plasmid, containing L1-RP coupled with an enhanced GFP retrotransposition reporter cassette or 99 RPS-GFP JM111 PUR plasmid, containing L1-RP with two-point mutations in ORF1 that abolish retrotransposition (L1-MT), was co-transfected together with 10 nM of siRNA (final concentration) using Lipofectamine 2000 (Invitrogen). All transfections were performed in triplicates. Five days post-transfection, the percent GFP-expressing cells was measured by flow cytometry (BD

Celesta or BD LSR Fortessa). Independent cells were transfected with enhanced GFP-N1-expressing plasmid to account for the effects of siRNA-mediated knockdown on transfection efficiency or cell proliferation/death.

5.4.6 Alu Retrotransposition Colony Formation Assay

Alu assay was performed as described in section 4.4.5 using both HeLa and ES2 cells.

5.4.7 Northern Blot

A protocol for Northern blot detection of L1 previously described (Deininger and Belancio, 2016) was followed with the following modifications. RNA from ES2 cells was treated with DNase I (Qiagen, 79254) and diluted in a loading buffer. RNA samples were heated at 70°C for 10 min and run on a 1% formaldehyde agarose gel. RNA was capillary transferred (Rio, 2015) to a Nylon membrane (Roche, 11417240001) overnight in 6x SSC buffer at room temperature. RNA was cross-linked to the membrane with 120 mJ UV using a Stratalinker 2400. The membrane was pre-hybridized for 30 min at 42 °C in 30% formamide, 1X Denhardt's solution, 1% SDS, 1mM NaCl, and 1µl/ml salmon sperm DNA and RNase free water.

Probes specific for L1HS were generated using two approaches. First, twelve 21-23 nt oligonucleotides were designed to recognize the 900 nt 5' end of L1HS specifically at regions absent from L1PA or with sequence differences compared to L1PA. As a loading control, 12 short oligonucleotides complementary to *GAPDH* mRNA were used. Probes recognizing L1HS or *GAPDH* were pooled and labeled with ^{32}Pγ-ATP as described (Llave et al., 2002); all sequences are listed in Table 5-4 below.

Table 5-4 Northern blot short probes used in Chapter 5 sequences.

Short Northern Probes	5`-3`sequence
L1-HS-5UTR-P1	CTAGTGAGATGAACCCGGTACC
L1-HS-5UTR-P2	CTTTGACTCGGAAAGGGAACTC
L1-HS-5UTR-P3	CGATTTTCCAGGTGCGTCCG
L1-HS-5UTR-P4	CGGCTGCTTTGTTTACCTAAGC
L1-HS-5UTR-P5	GGTCAGGGGTCAGGGACCCA
L1-HS-5UTR-P6	TGATGTACAGATGGGTTTTCGG
L1-HS-5UTR-P7	GGTGCCTCCCAGTTAGGCTG
L1-HS-5UTR-P8	TGTCAGTGTGCCCCTGCTGG
L1-HS-5UTR-P9	GCAATATTCGGGTGGGAGTGAC
L1-HS-5UTR-P10	GCCGTTTCTTAAGCCGGTCTGA
L1-HS-5UTR-P11	AGGTGTGGGATATAGTCTCGTG
L1-HS-5UTR-P12	TTGGAATACCCTGCCGTGTGAG
L1-PA-5UTR-P1	GAGTACCCGGCCGTGTGAGG
L1-PA-5UTR-P2	TGCGTCTAGTCGGCCATCTT
L1-PA-5UTR-P3	TGCTCTCGGTCCCGGGTCTG
GAPDH-P1	GAGAGAACAGTGAGCGCCTAGT
GAPDH-P2	ACCCGTTGACTCCGACCTTC
GAPDH-P3	TACGACTGCAAAGACCCGGAGC
GAPDH-P4	GGACCTCCATAAACCCACTTCT
GAPDH-P5	GCTACAGAAAGGTCAGCAGCTA
GAPDH-P6	GGCTTTGTGAGGAACTGGGAGA
GAPDH-P7	ACGAAGCCCTTCCAGGAGAAG
GAPDH-P8	CTCAGCACAAAGAACCCTCAGG
GAPDH-P9	CCAGGGCCGGCAGATGTCTTA
GAPDH-P10	GGACCTCCTGTTTCTGGGGACT
GAPDH-P11	GAAAGAAAGCGTCCCCCACCTA
GAPDH-P12	CAAGGCTGTTTGGAGGCCCTAC

As an alternative, L1HS was detected using a probe recognizing the 486 nt in the 5'UTR of L1HS. Probe DNA was amplified from HEK293T DNA by PCR using Taq DNA polymerase (BioBasic) using primer pairs 5'-gggaggaggagccaagat and 5'-ccggctgctttgtttaccta for L1HS, and 5'-accacagtccatgccatcac and 5'-gcttgacaaagtggtcgttg for GAPDH, respectively. PCR products were gel-purified using the Qiagen PCR cleanup kit. Radiolabeled probes were generated as described (DONG et al., 2013). Briefly, 25 ng of clean PCR product in 30 µl water with 125 ng random hexamer deoxynucleotide primers (IDT DNA) was heated for 2 min in a boiling water bath and snapped coldly on ice. dATP, dGTP and dTTP (Applied Biosystems, 250 µM final concentration) were added to 10 µl of 5x random priming buffer (250 mM Tris-Cl pH8.0, 25 mM MgCl2, 100 mM NaCl, 10 mM DTT, 1M HEPES pH 6.6), 5 µl [α-32P] dCTP (10 mCi/ml, 3000 Ci/nmol; Perkin Elmer), and the reaction brought to 50 µl with water. Five U of Klenow DNA polymerase (NEB) were added (60 min, room temperature), and the reaction was terminated with 10 µl stop buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM EDTA, 0.5% w/v SDS). Labeled probe heat-denatured before use in Northern blots. Probes were incubated with rotating membranes overnight at 42 °C. Membranes were washed twice in 5x SSC and once in 1x SSC.

5.4.8 Co-Immunoprecipitation and RNA Immunoprecipitation (RIP):

Cells were washed in cold PBS, lysed and scraped in RIPA lysis buffer or RIP lysis buffer. The cell lysate was incubated inverting for 20 min at 4 °C and centrifuged at 1000×g, 5 min to exclude insoluble materials. The supernatant was incubated inverting for 20 min at 4 °C with 10 µl of pre-washed protein G-Dynabeads to pre-clear. Beads were removed

using magnetic support (12321D, Thermo Fisher Scientific), and supernatant protein content was measured. Equal amounts of protein mass were incubated with equal amounts of the antibody of interest and its corresponding nonspecific IgG control antibody for 3-4 hr at 4 °C. 20 µl wet volume of protein G-Dynabeads were added to each IP for a further 1 hr at 4 °C. Beads were collected and washed with 200 µl RIPA buffer or RIP wash buffer three times at least. Co-Immunoprecipitation samples were processed for western blot analysis, and RIP samples were TRIzol (Invitrogen) RNA isolated as in the manufacturer's protocol.

5.4.9 ChIP and DRIP

Protocol for ChIP (Wiehle and Breiling, 2016) and DRIP (Sanz and Chédin, 2019) were followed with minor modifications described below.

5.4.9.1 ChIP

HeLa or ES2 cells were cultured on 100 mm dishes, and Fixation solution was added directly to them, followed by Glycine. Cells were scraped off, collected in tubes and centrifuged. Pellets were rewashed and resuspended in cell lysis buffer. Nuclei were collected by centrifugation and resuspended nuclear lysis buffer. Samples were sonicated immediately using Covaris S220 Ultra sonicator in Covaris snap cap microTUBE at 3 °C chilling system, 4-6 °C water, 175 peak power, 10% impact factor, 200 cycles, 560s timing. Sonicated suspensions were centrifuged, and supernatants were diluted four times in dilution buffer. Chromatin OD_{260} was measured after reverse cross-linking followed by DNA extraction, and at least 25µg of chromatin was used per IP. Supernatants were pre-cleared by pre-washed protein G-Dynabeads rotating for 20 min at 4 °C and were

incubated in 5µg of BRCA1 D-9 antibody or mouse mock IgG antibody for 3-4 hr at 4 °C. 20 µl wet volume of protein G-Dynabeads was added to each IP for 1 hr at 4 °C. Beads were collected and washed in 200 µl RIPA buffer six times at least and resuspended in 30µl in TE buffer to be used in qPCR after cross-links were reversed, and DNA was extracted using QIAquick PCR purification kit. The enrichment of L1 target regions was quantified. Three pairs of primers were designed based on the L1HS consensus sequence in Repbase at the element 5' end, middle, and at the 3' end. Oligos sequences are listed in Table 5-5 and the positive and negative controls used from (Hatchi et al., 2015).

5.4.9.2 DRIP

For each sample, 7-8 million cells were washed, trypsinized, neutralized with FBS containing media, and centrifuged. Cell pellets were washed and centrifuged again and resuspended in TE buffer. Samples were incubated in SDS (20%) and proteinase K (20mg/ml) to lyse overnight at 37°C. DNA was phenol/chloroform/isoamyl alcohol (25:24:1) extracted. Samples were centrifuged, and DNA supernatants were poured onto tubes containing 1/10 vol NaOAc (3M, pH5.2) and 2.5 vol of ethanol (100%, vol/vol) per sample. Precipitated DNA was spooled out and transferred to a new 2ml tube. DNA was washed in ethanol (80%, vol/vol), air-dried and resuspended in TE buffer and kept on ice for 1hr before adding restriction enzymes mix. The digestion mix was incubated overnight at 37°C until full digestion. Digested DNA was purified by phenol/chloroform/isoamyl alcohol, and air-dried DNA was resuspended in TE buffer. OD_{260} was measured, and 10µg of DNA was treated with RNase H (Thermo, EN0201) for 4h at 37°C followed by 5min incubation at 65°C. S9.6 antibody supernatant isolated from hybridoma cells

(ATCC® HB-8730™) was used for immunoprecipitation. 8µg DNA from RNase H treated or untreated samples were diluted in 500µl TE buffer. Protein G-Dynabeads were washed in 1X DRIP binding buffer three times and were added to 20µg S9.6 antibody supernatant and 10X DRIP binding buffer per IP. The antibody-beads mix was incubated overnight, rotating at 4°C. The supernatant was discarded, and beads were washed twice in 1X DRIP binding buffer. Diluted DNA in 10X DRIP binding buffer was added to the beads and incubated overnight at 4°C rotating. The supernatant was discarded, beads were washed similarly, and elution buffer with proteinase K was added. Bead tubes were parafilm sealed and incubated rotating in a hybridization oven at 55°C for 45min. Supernatants were transferred to new tubes, and DNA was Phenol/chloroform/isoamyl alcohol extracted. Air-dried DNA was resuspended in 50µl TE buffer and incubated on ice for 15min. Samples were run on qPCR against DRIP control targets listed in Table 5-5 below (Sanz and Chédin, 2019).

Table 5-5 ChIP and DRIP targets oligos used in Chapter 5 sequences

ChIP Oligos		
Oligos	**Forward**	**Reverse**
L1 Hs 5'	AGCCAAGATGGCCGAATAGG	AACCCGGTACCTCAGATGGA
L1 Hs mid	AAATCAGAGCAGAACTGAAGGAA	CCAGCTCCTGGATTCATTGAT
L1 Hs 3'	ACATGGACACAGGAAGGGGA	CTGCACCCACTAACGTGTCA
β-actin Intron	CGGGGTCTTTGTCTGAGC	CAGTTAGCGCCCAAAGGAC
β-actin Pause	GGGACT A TTTGGGGGTGTCT	TCCCATAGGTGAAGGCAAAG
DRIP Oligos		
Oligos	**Forward**	**Reverse**
RPL13A	AGGTGCCTTGCTCACAGAGT	GGTTGCATTGCCCTCATTAC
EGR1	GAACGTTCAGCCTCGTTCTC	GGAAGGTGGAAGGAAACACA

5.4.10 Pulse Labeling of RNA with Ethynyl-uridine

According to the manufacturer's instructions, pulse labeling was performed using the Nascent RNA Click-iT kit (Life Technologies). Briefly, cells were plated at 2.5×10^5 per well in a six-well plate and transfected with siRNA the following day. Seventy-two hours after transfection, ethynyl-uridine (0.2 mM) was added. Media was changed after 1 hr incubation, and RNA was extracted 1hr later using TRIzol (Invitrogen). RNA (5 µg) was labeled with biotin (0.5 mM Biotin Azide per sample). Pulse-labeled, biotinylated RNA (500 ng) was captured using Dynabeads Streptavidin T1 magnetic beads and used as a template for cDNA synthesis using MiScript II Reverse Transcriptase system (Qiagen).

5.4.11 Buffers

5.4.11.1 RIPA lysis buffer

10 mM Tris pH 7.4, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1% TritonX-100, 0.1% Sodium Deoxycholate, 0.1% SDS, Roche Complete Protease Inhibitor Cocktail Tablet

5.4.11.2 TBST

10 mM Tris-HCl pH 8, 150 mM NaCl, 0.05% Tween 20

5.4.11.3 RIP Lysis Buffer

50mM Tris-HCl pH 7.5, 150 mM NaCl, 1% NP40

5.4.11.4 RIP Wash Buffer

50mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% NP40

ChIP Fixation Solution

11% Formaldehyde (from a 37% stock equilibrated with methanol), 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 50 mM HEPES (pH 8), Roche Complete Protease Inhibitor Cocktail Tablet

5.4.11.5 ChIP Cell Lysis Buffer

5 mM PIPES (pH 8), 85 mM KCl, 0.5% NP40, Roche Complete Protease Inhibitor Cocktail Tablet

5.4.11.6 ChIP Nuclear Lysis Buffer

50 mM Tris–HCl (pH 8), 10 mM EDTA, 0.8% sodium dodecyl sulfate (SDS), Roche Complete Protease Inhibitor Cocktail Tablet

5.4.11.7 ChIP Dilution Buffer

10 mM Tris–HCl (pH 8.0), 0.5 mM EGTA, 1 % Triton X-100, 0.1 % Na-deoxycholate, 140 mM NaCl, Roche Complete Protease Inhibitor Cocktail Tablet

5.4.11.8 TE Buffer

1 mM EDTA, 10 mM Tris–HCl (pH 8.0)

10X DRIP Binding Buffer

100mM sodium phosphate, pH7, 1.4M NaCl, 0.5% Triton X-100, vol/vol

5.4.11.9 DRIP Elution Buffer

50mM Tris, pH8, 10mM EDTA, pH8, 0.5%SDS, vol/vol

6 CHAPTER 6: IMPACTS OF CHEMOTHERAPY USED IN CLINIC ON RETROTRANSPOSON EXPRESSION AND ACTIVITY

6.1 Background

The standard ovarian cancer patient care approach (as mentioned in chapter 1) includes debulking surgery and chemotherapy that consists of Carboplatin and Paclitaxel (Bell et al., 2011). If the surgery preceded the chemotherapy, the approach is adjuvant, and if the chemotherapy preceded the surgery, it is neoadjuvant. We discussed the differences between the two approaches in their application and impacts. Either approach is considered a primary therapy followed by maintenance therapy that may include the PARPi based on the case (Armstrong et al., 2021). BRCA1 status is significant in selecting treatment for both primary and maintenance therapy. High-grade serous ovarian cancer patients with BRCA1 deficiency (due to mutations or promoter hypermethylation) exhibited better overall survival and sensitivity to primary (platinum-based) chemotherapy than patients with *BRCA1* wild-type (Vencken et al., 2011; Yang et al., 2011; Rudaitis et al., 2014). Also, the maintenance therapy containing PARPi is more effective in patients with BRCA1 deficiency due to the synthetic lethality triggered in cancer cells (Farmer et al., 2005).

We demonstrated in chapter 5 evidence of the mechanism BRCA1 uses to regulate retrotransposon expression. Evidence demonstrated that treating cells with platinum-

based drugs such as Cisplatin or Carboplatin activates retrotransposon transcription (Rudin and Thompson, 2001; Guler et al., 2017). This activation of retrotransposon expression might be related to alleviating the epigenetic repression placed on their promotors (Guler et al., 2017). In addition, a subset of cancer cells could survive the lethal drug exposures by increasing the H3K9me3-mediated heterochromatin formation over L1 to restrict it (Guler et al., 2017). Alleviating this repression by HDACi treatment could activate L1 and sensitize these cells to lethal drug treatments (Guler et al., 2017). On the other hand, other evidence showed that out of eighty-eight tumor-specific L1 insertions identified in ovarian tumors, only one of these added a novel enhancer to the *STC18* gene that induced chemoresistance in cells having this mutation (Nguyen et al., 2018). These observations suggest that L1 activity may play a role in cancer cell sensitivity to chemotherapeutic drugs.

Literature on how PARP members are regulating retrotransposon is not clear. PARP1 interacts with ORF2 physically, and depletion of its endogenous levels reduced LINE-1 retrotransposition rates (Taylor et al., 2013). PARP2 promotes L1 retrotranspositions in conditions of genomic stress (Benitez-Guijarro et al., 2018; Miyoshi et al., 2019). Treating cells with Olaparib in retrotransposition assays was sufficient to decrease L1 canonical activity in TPRT proficient cells, not in those bearing EN mutation (Miyoshi et al., 2019). However, PARP13 (ZAP) restricts retrotransposons activity by preventing L1 mRNA accumulation in the cytoplasm (Moldovan and Moran, 2015). Colocalization of ORF1p and L1 RNA in SGs possibly promoted the recruitment of RNA degradation machinery (Moldovan and Moran, 2015). Some of the contradictory roles of PARP members in

retrotransposon regulation need further investigation. That includes an evaluation of retrotransposon RNA expression upon chemotherapeutic use and PARP inhibition.

This chapter wanted to examine the effect of both primary and maintenance chemotherapy exposures on retrotransposon expression. This investigation will increase our understanding of the mechanisms that might be involved in retrotransposon regulation.

6.2 Results

6.2.1 No effect of Olaparib treatment on levels of retrotransposon RNA expression or genomic insertions in cell models

The previous chapter demonstrated how BRCA1 regulates retrotransposon transcription independent of its roles in controlling DNA damage repair and R-loop resolution. Chapter 4 showed evidence of mitochondrial regulation of retrotransposon expression (Figure 4-3). PARPs are NAD+ consumers, as mentioned above, and NAD+ homeostasis is vital for cellular metabolic activities taking place in the mitochondria (Katsyuba et al., 2020). Also, PARP inhibitors are used as maintenance chemotherapy to treat ovarian cancer patients in the clinic. Therefore, we sought to test whether inhibiting PARP will impact levels of retrotransposon expression and activity. We observed no significant difference in L1 and *Alu* RNA expression between DMSO vehicle control and 10μM Olaparib treated ES2 and MDA-MB-231 Cells (Figure 6-1A). Both cell lines express wild-type *BRCA1*, but they express variants of it. The *BRCA1* variants of ES2 cells were mentioned previously (chapter 5). The triple-negative (that lack expression of ER, progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) genes) MDA-MB-231 cells

express two *BRCA1* variants 561-34C > T and 1186A > G (Elstrodt et al., 2006). As expected from the literature, Olaparib treatment reduced levels of genomic insertion in cells compared to their counterparts (Figure 6-1 B,C). This evidence suggests that the observed effect of Olaparib on cancer cells is independent of retrotransposon regulation.

6.2.2 Adjuvant chemotherapy reduces levels of retrotransposon RNA expression

Next, we wanted to examine the effect of first-line chemotherapy as single agents or in combination on levels of retrotransposon RNA. Although treating ES2 cells with Carboplatin increased L1 and *AluY5a* RNA expression significantly, Taxol treatment for 24hr did not change levels of retrotransposon RNA (Figure 6-2 A,B). Interestingly, when both drugs were combined to treat ES2 cells for the same duration (24hr), levels of retrotransposon RNA decreased significantly (Figure 6-2 C). We performed a secondary analysis to compare the effect of different chemotherapeutic regimens to treat high-grade serous ovarian cancer patients based on TCGA data. From the hundreds of patients in TCGA, only some had recorded data that included the chemotherapy protocol followed. The patient information indicated that forty-seven patients received neoadjuvant therapy (involved neoadjuvant and interval surgery followed by postoperative chemotherapy).

On the other hand, twenty-eight patients received adjuvant therapy (involved primary surgery followed by postoperative chemotherapy). The first group demonstrated markedly higher *Alu* and L1PA RNA expression than the second group of patients (Figure 6-2 D). Intriguingly, the effect on retrotransposon RNA expression using the combination of Carboplatin and Taxol treatment in cell culture mimicked the effect of using these drugs

in patients (Figure 6-2 C,D). This observation suggests that retrotransposon expression may change by changing chemotherapeutic drugs used and the timing of their administration.

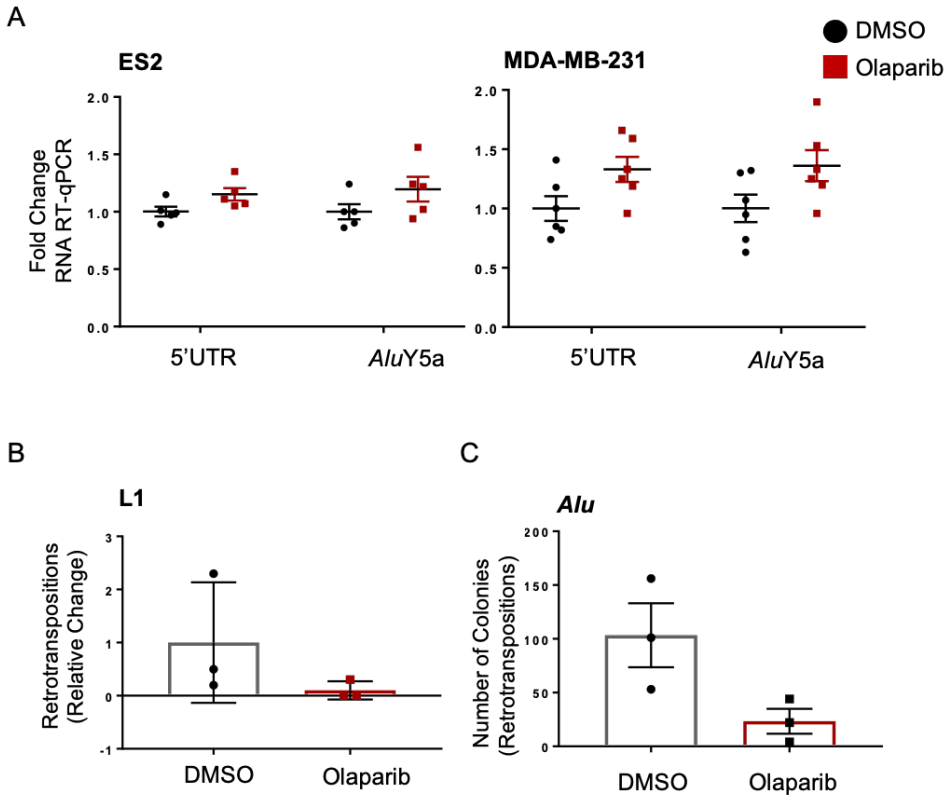


Figure 6-1 Effect of PARPi (Olaparib) on retrotransposon RNA expression and activity
A. RT-qPCR of LINE-1 (5'UTR) or *AluY*5a RNA after 24hr treatment of 10μM Olaparib treatment. ES2 cells (left panel) and MDA-MB-231 cells (right panel), N=5 and N=6 independent biological replicates respectively. **B,C.** Relative rate of LINE-1 (**B**) or Alu retrotranspositions (**C**) in cells after treatment with vehicle control or 10μM Olaparib. Horizontal lines represent averages and error bars indicate standard error of the mean. Two-way ANOVA with Sidak correction (**A**) and student t-test (**B,C**) were used for analysis.

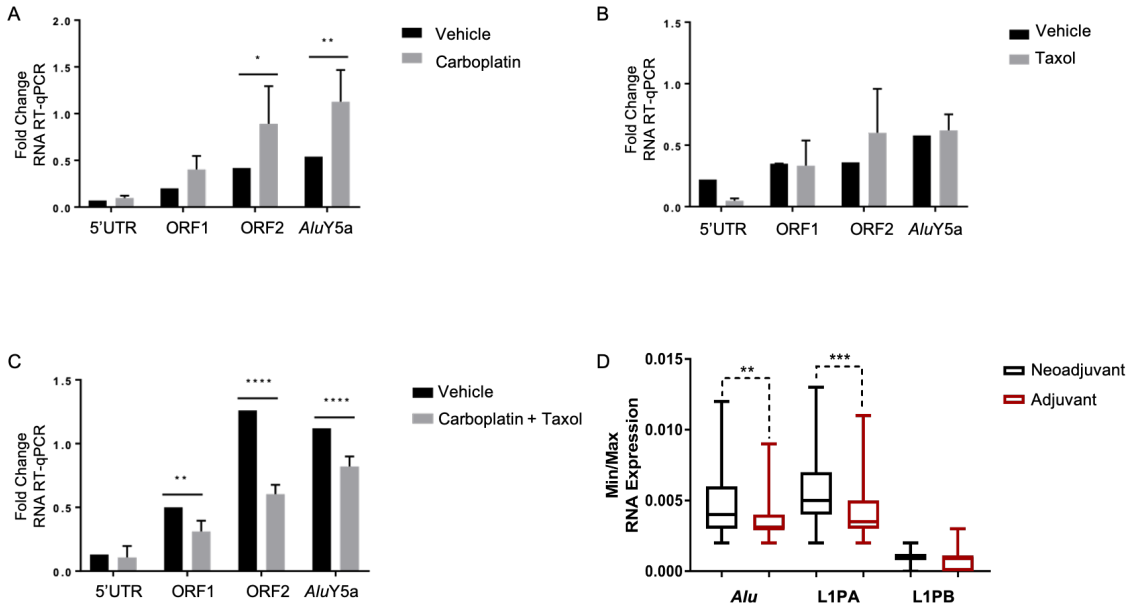


Figure 6-2 Effect of platinum–taxane chemotherapy on retrotransposon RNA expression
A-C. RT-qPCR of LINE-1 and *AluY5a* RNA after 24hr treatment of 50µM Carboplatin (**A**), 2.5nM Paclitaxel (**B**), or a combination of both (**C**) treatments. **D.** Group comparison analysis between retrotransposon RNA expression in HGSOV patients received neoadjuvant and those who received adjuvant first-line chemotherapy options.

6.3 Discussion

The response to chemotherapy treatment can involve several factors, including the genetic makeup of the patient, ovarian cancer subtype, and the timing of primary chemotherapy provided. Parts of the patient genetic background are related to HR pathway proficiency, including BRCA1 status. Retrotransposon expression could also play a role in the cellular sensitivity to specific chemotherapeutic choices. Since one of the main regulatory mechanisms revealed in this thesis is related to BRCA1, PARP inhibition was a prominent chemotherapy option we sought to examine.

The effect of Olaparib on cancer cells is well established, and it is one of the therapeutic maintenance choices in the current management of ovarian cancer (Banerjee and Lord, 2020). We noticed an insignificant increase in retrotransposons expression upon PARP

inhibition. This mild increase could be related to the known restriction effect of PARP13 on retrotransposons (Moldovan and Moran, 2015). We also observed substantial cell death in cells treated with BRCA1 siRNA and Olaparib (data not shown). This observation agrees with the synthetic lethality previously demonstrated when PARP is inhibited in BRCA1 deficient cells (Farmer et al., 2005). In general, the presented data suggests that the effect of PARP inhibition on cancer cell survival is more pronounced than the effect on retrotransposon RNA expression or insertion levels. This evidence also indirectly supports our earlier data that BRCA1 regulates retrotransposon expression by a mechanism independent of DNA damage repair pathways, such as those regulated by PARPs.

Here our data indicates that the patient treatment regimen could change the retrotransposon expression. One of the limitations of this analysis is the knowledge of BRCA1 status in patients of the compared groups (Figure 6-2 D). Survival analysis previously showed that patients with epigenetically silenced *BRCA1* had poorer outcomes than patients with wild-type *BRCA1* and mutated *BRCA1* (Bell et al., 2011). Patients with *BRCA1* mutations had the best prognosis than the other groups (Bell et al., 2011). This difference in the patient prognosis could be related to the chemotherapy used and the pathways activated by this treatment not directly related to the genetic silencing. Other evidence mentioned earlier compared the outcomes of patients with BRCA1 dysfunction to patients with *BRCA1* wild-type outcomes after receiving chemotherapy. The patients who exhibited BRCA1 dysfunction (including epigenetic silencing) survived better than the others (Vencken et al., 2011; Yang et al., 2011; Rudaitis et al., 2014). This evidence

may imply that the chemotherapeutic treatments activate signaling pathways that contribute to the improved survival rates of patients. These pathways could involve retrotransposon activation, which triggers an immune response and promotes cancer cell death. This assumption is supported by the evidence that treating cells with Carboplatin activated the canonical and non-canonical STING signaling pathways (Zhou et al., 2021). This activation augmented CD8+ T-cell infiltration, increased PD-L1 expression, and potentiated the anti-tumor effect of PD-1 inhibitors (Zhou et al., 2021). It is possible that upstream of this mechanism, Carboplatin-mediated retrotransposon activation was the trigger of cGAS-STING. This hypothesis could be tested in future investigations and may suggest another strategy to enhance the effect of cancer immunotherapy.

Although limited by not testing for BRCA1 depletion, our evidence can suggest roles of retrotransposon activation in the response of chemotherapeutic therapy of ovarian cancer. Future experiments may include BRCA1 siRNA downregulated samples, samples treated with RTi, and models to test for chemoresistance. RTi treatment could serve as a reasonable negative control to examine if the effect is related to the retrotransposon activity. Also, incorporating different types of RTi into the experiment might suggest a specific mechanism to investigate. Modulating the retrotransposon levels may alter the cellular response to chemotherapy treatments and add knowledge to the field of chemoresistance.

6.4 Methods

6.4.1 RT-qPCR, L1, and Alu Retrotransposition Assays

Assays were performed as described in the previous chapters with the addition of drug treatments. 10μM Olaparib (Cedarlane, A10111), 2.5nM Paclitaxel (Taxol) (Selleckchem, S1150), and 50μM Carboplatin (Sigma, C 2538).

7 CHAPTER 7: GENERAL DISCUSSION

Completing the human genome project initiated the transformation of 'junk DNA' into a relevant subject of study. The development of cell culture and bioinformatics tools to study them paved the way to understand TEs evolution, biology and activity. One of the questions that attracted many scientists in the last decades has been: "How does retrotransposon activity affect cancer?" Ongoing research and significant observations contributed to an understanding of the regulation and impact of retrotransposons in various cancer types.

High L1 RNA expression predicted improved outcomes in high-grade serous ovarian but not basal-like breast cancer patients. Other reports showed that overexpressing retrotransposon caused immunogenic responses and enhanced survival in cancer patients and animal models (Chiappinelli et al., 2015; Guler et al., 2017; Natoli et al., 2021). Our data suggest distinct associations between retrotransposons and IFN signaling in breast (strong correlations) and ovarian tumors (weak correlations). This distinct correlation among different cancer types could be related to tumor-specific characteristics that may alter the tumor microenvironment and immune responses accordingly.

Previous evidence demonstrated that tumors with *TP53* mutations had low immune activity and higher loads of L1 insertions than tumors with wild-type *TP53* (Jung et al., 2018). The weak immune signature in ovarian tumors could be related to *TP53* mutations occurring in most high-grade serous ovarian tumors (Bell et al., 2011; Chien et al., 2015). This assumption is consistent with previous reports indicating that *TP53* has

immunomodulatory roles and its dysfunction associates with immunosuppression (Rooney et al., 2015; Cui and Guo, 2016). Evidence from colon cancer shows that in response to viral infection in cells, p53 induced IFN-dependant antiviral response by activating IFN-stimulated genes (ISGs) (Muñoz-Fontela et al., 2008). Another piece of evidence showed that p53 cooperates with DNA methylation to maintain the silencing of SINEs and other non-coding RNAs (Leonova et al., 2013). The p53-deficient cells in this study exhibited high expression of SINE elements accompanied by high type I IFN response (Leonova et al., 2013). However, not all tumors exhibit the same type of *TP53* mutation, and not all mutations result in p53 protein deficiency (Goldstein et al., 2011). *TP53* mutation can contribute to tumorigenesis by the loss of p53 function and by a gain of mutant functions (Goldstein et al., 2011). The majority of *TP53* mutations in high-grade serous ovarian tumors are gain-of-function mutations, which means p53 is expressed but with altered functions (Cole et al., 2016). On the other hand, most *TP53* mutations in basal-like breast tumors are loss-of-function mutations (cBioPortal).With that in mind, p53 may play a significant role in the heterogenous immune responses noticed among cancers in our study and other studies. This change in the immune signature could impact the retrotransposon expression and activity in tumors. *TP53* gain of function mutations in high-grade serous ovarian tumors could reduce retrotransposon expression and associated IFN response. In contrast, frequent *TP53* loss of function mutations in basal-like breast cancer could increase retrotransposon expression and associated IFN response.

Reprogramming energy metabolism is a cancer hallmark, which helps cancer cells to manage their energy resources and sustain their survival advantage (Hanahan and

Weinberg, 2011). Evidence suggests the involvement of p53 in cellular metabolism modulating the utilization of oxidative phosphorylation or glycolytic pathways in cancer (Matoba et al., 2006). The cancer cells tested here (ES2 and MDA -MB-231) were depending on oxidative phosphorylation, not on glycolysis to manage their energy needs (Seahorse assay data), unlike other cancer cells, which are primarily glycolytic (Hanahan and Weinberg, 2011). This effect could be related to the high nutrient content of culture media the cells are grown in whereas cancer cells in tumor are usually deprived of nutrients. *TP53* mutation in these cells could also play a role in altering their mitochondrial homeostasis. In addition, BRCA1 exerts an antioxidant effect in oxidative stress conditions by interacting with p53 and inducing p21 expression (Chai et al., 1999). p21, in turn, interacts with NRF2 (Nuclear factor erythroid 2-related factor 2), a transcription factor involved in inducing transcription of antioxidant enzymes and regulated by KEAP1 (Kelch Like ECH Associated Protein 1)- mediated ubiquitination (Chen et al., 2009; Mitsuishi et al., 2012). It is possible that in p53 mutated high-grade serous ovarian tumors, this antioxidant role of BRCA1 is disturbed, which can alter the mitochondrial activity. This alteration and the BRCA1-related roles may contribute to the elevated levels of retrotransposon RNA expression in these cells when treated with FCCP. Both FCCP and BAM15 induce a state of maximal respiratory capacity (Park et al., 2002; To et al., 2010; Alexopoulos et al., 2020) that resembles what happens in rapidly dividing cells like stem cells where ATP demand is high to support the cellular anabolic needs.

The mitochondrial regulation of retrotransposon expression and possibly activity could play a role in other pathologies such as Bipolar Disorder (BD). This disorder is

characterized by presenting two phases of energy availability according to the phase of illness (mania versus depression) (Morris et al., 2017). Mitochondrial activity increases in mania and decreases in depression (Morris et al., 2017). Clinical reports indicate the involvement of inflammatory responses in BD pathology, and patients are prone to develop autoimmune disorders such as hyperthyroidism and rheumatoid arthritis (Fries et al., 2019). High levels of circulating DNA segments of unspecified nuclear origin are detected in serum extracted from patients during acute episodes (Stertz et al., 2015). BD patients show activated TLRs and NF-κB pathways consistent with the trigger imposed by circulating DNA on these receptors (Fries et al., 2019). Brain samples of thirty-eight BD patients showed high HERV-K(HML2) expression in a study investigating TE roles in mood disorders (Frank et al., 2005). Another evidence from peripheral blood extracted from nighty nine BD patients and 92 healthy subjects observed increased levels of hypomethylation in the L1 CpGs of patients compared to the controls (Li et al., 2018). Although at what BD phase the blood was taken is not included, this evidence suggests the involvement of retrotransposons in the pathology of BD. Knowledge about what triggered this TE activation, how and at which phase of the illness would be interesting to investigate further. From the findings in this thesis, we can speculate that the increase in mitochondrial activity in the mania phase could trigger retrotransposon expression through a mechanism (yet to be explored) related to epigenetic regulation.

Retrotransposon activity is also implicated in multiple neurodegenerative disorders such as AGS, Multiple Sclerosis, Amyotrophic Lateral Sclerosis, and Alzheimer's Disease (Tam et al., 2019). Evidence indicates the detection of retrotransposon-associated RNAs, DNAs

and proteins in patient samples in these diseases (Tam et al., 2019). AGS and Multiple Sclerosis disorders show evidence of retrotransposon-triggered inflammatory responses. While Amyotrophic Lateral Sclerosis and Alzheimer's Disease show neurotoxic impacts of retrotransposon activity (Tam et al., 2019). Mitochondrial dysfunction is a shared characteristic among these disorders, which might play a role in the disturbance of retrotransposon activity in patients (Burté et al., 2015). Applying a similar strategy to identify factors regulating and regulated by retrotransposons in studying these disorders may enhance our understanding of retrotransposon activity triggers and suggest ways to target them for the benefit of patients.

BRCA1 is a tumor suppressor; however, during tumorigenesis, pathogenic or likely pathogenic variants of *BRCA1* can emerge to function differently (Sayaman et al., 2021). In triple-negative breast tumors, these *BRCA1* variants can act as immune response modulators (Sayaman et al., 2021). Since tumors with BRCA1 mutations are characterized by high immune infiltration compared to other breast cancer subtypes (Hendrickx et al., 2017), BRCA1 may exert some tissue-restricted effects that modulate tumor-derived immune responses (Sayaman et al., 2021). Triple-negative breast tumors and high-grade serous ovarian tumors share some characteristics, such as the high frequency of *TP53* mutations, mutations in *NF1* and *RB1* at lower frequencies, and *BRCA1* or *BRCA2* inactivation (Ciriello et al., 2012; Koboldt et al., 2012). In addition to the mutually exclusive loss of function in similar signaling pathways (Ciriello et al., 2012; Koboldt et al., 2012). Both cancer subtypes also co-existed in the same patients in a higher number than expected in a study compared the co-occurrence of ovarian and breast tumors of all

subtypes (Begg et al., 2017). This observation indicates a possible common etiology is shared by both ovarian and breast cancer subtypes (Begg et al., 2017). Our analyses noticed a distinct phenotype between ovarian and breast tumors tested in control of retrotransposon expression. This observation can be related to other factors in the tumor microenvironment, activating different signaling pathways, including immune responses. Although the restriction of our analysis to specific subsets of these cancers is a limiting factor, the regulation of retrotransposons by BRCA1 may be shared and justify the investigation in further tumor subtypes.

Retrotransposon expression is not the sole factor predicting survival in ovarian cancer patients in our analyses. Retrotransposons act within a whole system to regulate numerous RNAs operating in different pathways, or these RNAs regulate levels of retrotransposon RNAs, which can contribute to the retrotransposon activity associated with tumor heterogeneity (Nguyen et al., 2018). Tumor heterogeneity can influence the response of patients to chemotherapy (Dagogo-Jack and Shaw, 2018). Although the combination of Carboplatin and Paclitaxel is the standard option for primary care of ovarian cancer patients, clinical trials are not in agreement on the best time to initiate them (Kurnit et al., 2021). Evidence provided here indicates that drug choices and timing can alter retrotransposon expression, which may, in turn, play a role in predicting patient outcomes. However, these hypotheses call for further investigation.

8 CHAPTER 8: GENERAL CONCLUSION

The work presented in this thesis uncovered significant aspects of retrotransposon regulation and activity in cancer. We found that tumor-specific retrotransposon insertions impose minimal impacts on high-grade serous ovarian cancer tumor development. By a pan-genomic study, we identified factors and pathways involved in retrotransposon expression regulation. Inflammatory responses and mitochondrial activity were among these pathways. This evidence underscores the possible impact of retrotransposons in other conditions where these pathways are interrupted. In addition, this study exposed a mechanism by which BRCA1 regulates L1 transcription by binding its genomic DNA. These findings support the hypothesis that in untransformed cells, normal levels of BRCA1 may maintain a threshold of retrotransposon accumulation by inducing type I IFN response and possibly apoptosis. In ovarian cancer, retrotransposon expression regulates and is regulated by a complex network of correlating RNAs that may play a role in predicting survival in patients. The followed genome-wide strategy can help identify the causes and consequences of retrotransposon expression in other cancer types and other disorders.

APPENDICES

1. Appendix A

Supplementary Table 1 includes sheets from A-C can be downloaded from supplementary data section in the published paper Alkailani *et al*, A genome-wide strategy to identify causes and consequences of retrotransposon expression finds activation by BRCA1 in ovarian cancer, *NAR Cancer*, Volume 3, Issue 1, March 2021, zcaa040, https://doi.org/10.1093/narcan/zcaa040 or by following the link below:

https://oup.silverchair-cdn.com/oup/backfile/Content_public/Journal/narcancer/3/1/10.1093_narcan_zcaa040/1/zcaa040_supplemental_files.zip?Expires=1618510313&Signature=noP3IRNZleLOJlBpMl1JIFx6DGaowmB1SXsNm~6WSJk0qgXVZhbrFZNQESXX25u7dMNWoWti~6tAcatWyQvqA9lLGPbVrNjUu9szP2kSYiAZhPJdovl3KNqJGnaQjBy0xPKHlQpWp-ZNi3c8M5UNjhLTdeb3VWyzg4Dlz7F0Xrf-VHep-AAZhSSEZtpMSK1NhgL5wb33XmHdeRmOQR5C8~yDxSvD9immKA1T8NPLG1v4i9wFsEw9irooPHgXirky2r5QebqJvAzWQzpkusKLaXBPtkGxalq2phEkTU8ntj7pF~ysHkdzakb8qq2JvHJaxAyboVxxwyUV5uXQ6kbIag__&Key-Pair-Id=APKAIE5G5CRDK6RD3PGA

2. Appendix B

Dataset 1 can be downloaded from supplementary data section in the published paper Alkailani *et al*, A genome-wide strategy to identify causes and consequences of retrotransposon expression finds activation by BRCA1 in ovarian cancer, *NAR Cancer*, Volume 3, Issue 1, March 2021, zcaa040, https://doi.org/10.1093/narcan/zcaa040 or by following the link below:

https://oup.silverchair-cdn.com/oup/backfile/Content_public/Journal/narcancer/3/1/10.1093_narcan_zcaa040/1/zcaa040_supplemental_files.zip?Expires=1618510313&Signature=noP3IRNZleLOJlBpMl1JIFx6DGaowmB1SXsNm~6WSJk0qgXVZhbrFZNQESXX25u7dMNWoWti~6tAcatWyQvqA9lLGPbVrNjUu9szP2kSYiAZhPJdovl3KNqJGnaQjBy0xPKHlQpWp-ZNi3c8M5UNjhLTdeb3VWyzg4Dlz7F0Xrf-VHep-AAZhSSEZtpMSK1NhgL5wb33XmHdeRmOQR5C8~yDxSvD9immKA1T8NPLG1v4i9wFsEw9irooPHgXirky2r5QebqJvAzWQzpkusKLaXBPtkGxalq2phEkTU8ntj7pF~ysHkdzakb8qq2JvHJaxAyboVxxwyUV5uXQ6kbIag__&Key-Pair-Id=APKAIE5G5CRDK6RD3PGA

3. Appendix C

Dataset 2 can be downloaded from supplementary data section in the published paper Alkailani *et al*, A genome-wide strategy to identify causes and consequences of retrotransposon expression finds activation by BRCA1 in ovarian cancer, *NAR Cancer*, Volume 3, Issue 1, March 2021, zcaa040, https://doi.org/10.1093/narcan/zcaa040 or by following the link below:

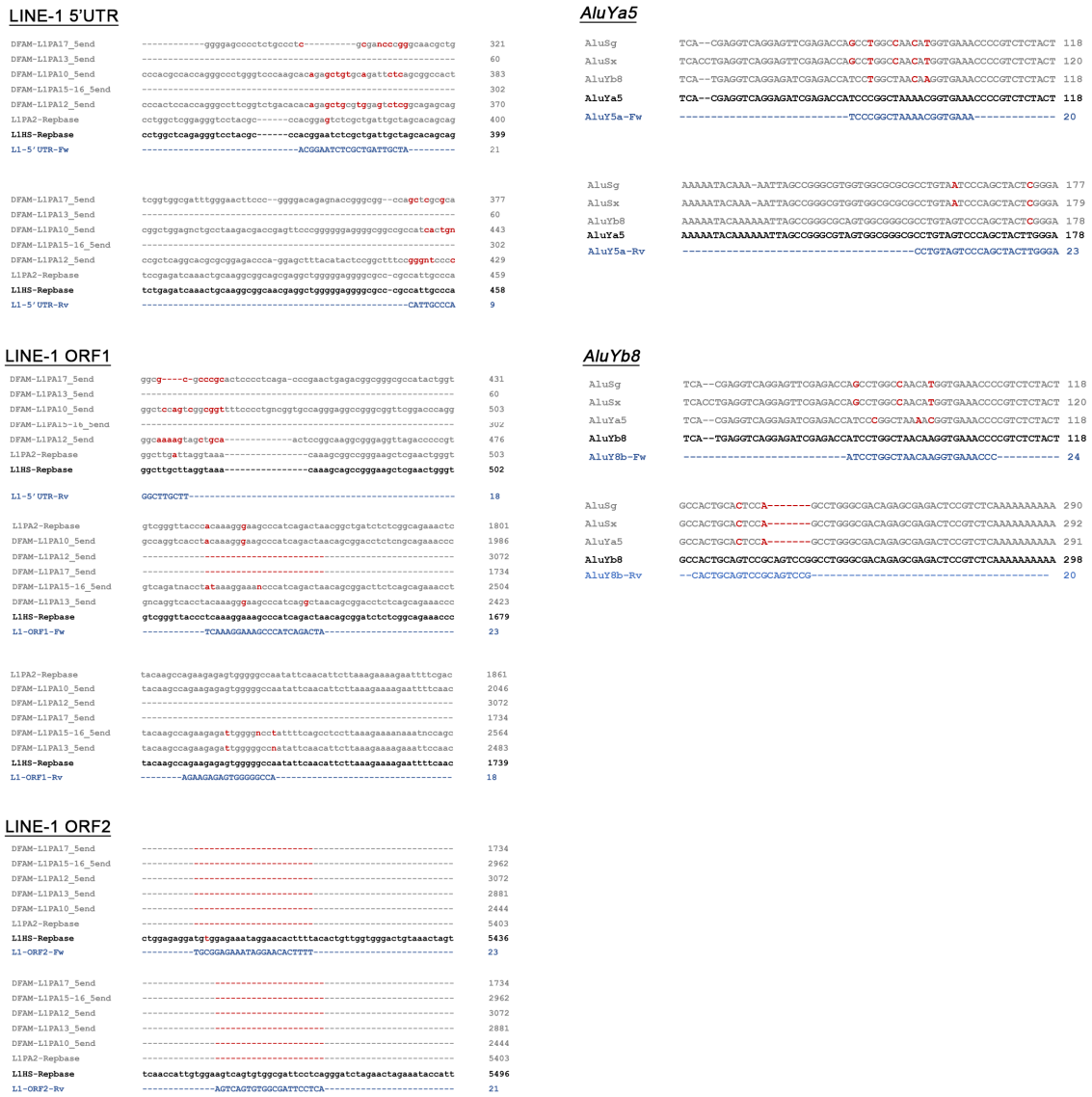
https://oup.silverchair-cdn.com/oup/backfile/Content_public/Journal/narcancer/3/1/10.1093_narcan_zcaa040/1/zcaa040_supplemental_files.zip?Expires=1618510313&Signature=noP3IRNZleLOJlBpMl1JIFx6DGaowmB1SXsNm~6WSJk0qgXVZhbrFZNQESXX25u7dMNWoWti~6tAcatWyQvqA9lLGPbVrNjUu9szP2kSYiAZhPJdovl3KNqJGnaQjBy0xPKHlQpWp-ZNi3c8M5UNjhLTdeb3VWyzg4Dlz7F0Xrf-VHep-AAZhSSEZtpMSK1NhgL5wb33XmHdeRmOQR5C8~yDxSvD9immKA1T8NPLG1v4i9wFsEw9irooPHgXirky2r5QebqJvAzWQzpkusKLaXBPtkGxalq2phEkTU8ntj7pF~ysHkdzakb8qq2JvHJaxAyboVxxwyUV5uXQ6kbIag__&Key-Pair-Id=APKAIE5G5CRDK6RD3PGA

4. Appendix D

Supplementary Table 2, includes sheets from A-G can be downloaded from supplementary data section in the published paper Alkailani *et al*, A genome-wide strategy to identify causes and consequences of retrotransposon expression finds activation by BRCA1 in ovarian cancer, *NAR Cancer*, Volume 3, Issue 1, March 2021, zcaa040, https://doi.org/10.1093/narcan/zcaa040 or by following the link below:

https://oup.silverchair-cdn.com/oup/backfile/Content_public/Journal/narcancer/3/1/10.1093_narcan_zcaa040/1/zcaa040_supplemental_files.zip?Expires=1618510313&Signature=noP3IRNZleLOJlBpMl1JIFx6DGaowmB1SXsNm~6WSJk0qgXVZhbrFZNQESXX25u7dMNWoWti~6tAcatWyQvqA9lLGPbVrNjUu9szP2kSYiAZhPJdovl3KNqJGnaQjBy0xPKHlQpWp-ZNi3c8M5UNjhLTdeb3VWyzg4Dlz7F0Xrf-VHep-AAZhSSEZtpMSK1NhgL5wb33XmHdeRmOQR5C8~yDxSvD9immKA1T8NPLG1v4i9wFsEw9irooPHgXirky2r5QebqJvAzWQzpkusKLaXBPtkGxalq2phEkTU8ntj7pF~ysHkdzakb8qq2JvHJaxAyboVxxwyUV5uXQ6kbIag__&Key-Pair-Id=APKAIE5G5CRDK6RD3PGA

5. Appendix E



Alignment of Primers for L1HS, *AluYa5*, and *AluYb8* with older variants of L1 and *Alu* elements. Forward (Fw) and reverse (Rv) primers for each primer pair are shown in blue. Sequence changes which distinguish L1HS, *AluYa5*, and *AluYb8* from their older counterparts are shown with red text. Alignments were generated with Clustal Omega.

6. Appendix F

OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS

Feb 05, 2021

This Agreement between Maisa Alkailani ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

License Number	5002611244482
License date	Feb 05, 2021
Licensed content publisher	Oxford University Press
Licensed content publication	NAR Cancer
Licensed content title	A genome-wide strategy to identify causes and consequences of retrotransposon expression finds activation by BRCA1 in ovarian cancer
Licensed content author	Alkailani, Maisa; Palidwor, Gareth
Licensed content date	Jan 6, 2021
Type of Use	Thesis/Dissertation

OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS

Feb 05, 2021

This Agreement between Maisa Alkailani ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

License Number	5002620366737
License date	Feb 05, 2021
Licensed content publisher	Oxford University Press
Licensed content publication	NAR Cancer
Licensed content title	A genome-wide strategy to identify causes and consequences of retrotransposon expression finds activation by BRCA1 in ovarian cancer
Licensed content author	Alkailani, Maisa; Palidwor, Gareth
Licensed content date	Jan 6, 2021
Type of Use	Thesis/Dissertation
Institution name	
Title of your work	Doctor of Philosophy Thesis
Publisher of your work	University of Ottawa
Expected publication date	Apr 2021

OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS

Feb 05, 2021

This Agreement between Maisa Alkailani ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

License Number	5002621161678
License date	Feb 05, 2021
Licensed	

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

May 17, 2021

This Agreement between Maisa Alkailani ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5071521089348
License date	May 17, 2021
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Reviews Cancer
Licensed Content Title	Transposable elements in cancer
Licensed Content Author	Kathleen H. Burns
Licensed Content Date	Jun 9, 2017
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic
Portion	figures/tables/illustrations

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

May 20, 2021

This Agreement between Maisa Alkailani ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	5073120506815
License date	May 20, 2021
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Nature Reviews Genetics
Licensed Content Title	A change of view: homologous recombination at single-molecule resolution
Licensed Content Author	Kyle Kaniecki et al
Licensed Content Date	Dec 11, 2017
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	print and electronic

BIBLIOGRAPHY

Akaike, H. (1974). A New Look at the Statistical Model Identification. IEEE Trans. Automat. Contr. *19*, 716–723.

Alexopoulos, S.J. et al. (2020). Mitochondrial uncoupler BAM15 reverses diet-induced obesity and insulin resistance in mice. Nat. Commun. *11*, 2397.

Alkailani, M. et al. (2021). A genome-wide strategy to identify causes and consequences of retrotransposon expression finds activation by BRCA1 in ovarian cancer. NAR Cancer *3*.

Alpatov, R. et al. (2004). Nuclear Speckle-Associated Protein Pnn/DRS Binds to the Transcriptional Corepressor CtBP and Relieves CtBP- Mediated Repression of the E-Cadherin Gene. Mol. Cell. Biol. *24*, 10223–10235.

Anderson, S.F. et al. (1998). BRCA1 protein is linked to the RNA polymerase II holoenzyme complex via RNA helicase A. Nat Genet *19*, 254–256.

Armstrong, D.K. et al. (2021). Ovarian cancer, version 2.2020. JNCCN J. Natl. Compr. Cancer Netw. *19*, 191–226.

Aschacher, T. et al. (2016). LINE-1 induces hTERT and ensures telomere maintenance in tumour cell lines. Oncogene *35*, 94–104.

Ashburner, M. et al. (2000). Gene ontology: Tool for the unification of biology. Nat. Genet. *25*, 25–29.

Ataseven, B. et al. (2016). Prognostic impact of debulking surgery and residual tumor in patients with epithelial ovarian cancer FIGO stage IV. Gynecol. Oncol. *140*, 215–220.

Athanikar, J.N. et al. (2004). A YY1-binding site is required for accurate human LINE-1

transcription initiation. Nucleic Acids Res. *32*, 3846–3855.

Auersperg, N. (2013). Ovarian surface epithelium as a source of ovarian cancers: Unwarranted speculation or evidence-based hypothesis? Gynecol. Oncol. *130*, 246–251.

Babaian, A., and Mager, D.L. (2016). Endogenous retroviral promoter exaptation in human cancer. Mob. DNA *7*, 1–21.

Babushok, D. V., and Kazazian, H.H. (2007). Progress in understanding the biology of the human mutagen LINE-1. Hum. Mutat. *28*, 527–539.

Baker, M. (2010). Epigenome: Mapping in motion. Nat. Methods *7*, 181–186.

Balaj, L. et al. (2011). Tumour microvesicles contain retrotransposon elements and amplified oncogene sequences. Nat. Commun. *2*, 180.

Banerjee, S.N., and Lord, C.J. (2020). First-line PARP inhibition in ovarian cancer — standard of care for all? Nat. Rev. Clin. Oncol. *17*, 136–137.

Bao, W. et al. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob. DNA *6*, 11.

Bau, D.T. et al. (2004). Breast cancer risk and the DNA double-strand break end-joining capacity of nonhomologous end-joining genes are affected by BRCA1. Cancer Res *64*, 5013–5019.

Becker, K.G. et al. (1993). Binding of the ubiquitous nuclear transcription factor YY1 to a cis regulatory sequence in the human LINE-1 transposable element. Hum. Mol. Genet. *2*, 1697–1702.

Bedell, J.A. et al. (2000). MaskerAid: A performance enhancement to RepeatMasker. Bioinformatics *16*, 1040–1041.

Begg, C.B. et al. (2017). Examining the common aetiology of serous ovarian cancers and

basal-like breast cancers using double primaries. Br. J. Cancer *116*, 1088–1091.

Belancio, V.P. et al. (2010). Somatic expression of LINE-1 elements in human tissues. Nucleic Acids Res. *38*, 3909–3922.

Bell, D. et al. (2011). Integrated genomic analyses of ovarian carcinoma. Nature *474*, 609–615.

Bell, D.A. (2005). Origins and molecular pathology of ovarian cancer. Mod. Pathol. *18*, 19–32.

Benecke, H. et al. (2005). The U11/U12 snRNP 65K protein acts as a molecular bridge, binding the U12 snRNA and U11-59K protein. EMBO J. *24*, 3057–3069.

Benezra, M. et al. (2003). BRCA1 augments transcription by the NF-κB transcription factor by binding to the rel domain of the p65/RelA subunit. J. Biol. Chem. *278*, 26333–26341.

Benitez-Guijarro, M. et al. (2018). RNase H2, mutated in Aicardi-Goutières syndrome, promotes LINE-1 retrotransposition. EMBO J. *37*.

Boeke, J.D. et al. (1985). Ty elements transpose through an RNA intermediate. Cell *40*, 491–500.

du Bois, A. et al. (2003). A randomized clinical trial of cisplatin/paclitaxel versus carboplatin/paclitaxel as first-line treatment of ovarian cancer. J. Natl. Cancer Inst. *95*, 1320–1330.

Boissinot, S., and Furano, A. V. (2001). Adaptive evolution in LINE-1 retrotransposons. Mol. Biol. Evol. *18*, 2186–2194.

Bourque, G. et al. (2018). Ten things you should know about transposable elements 06 Biological Sciences 0604 Genetics. Genome Biol. *19*, 199.

Brégnard, C. et al. (2016). Upregulated LINE-1 Activity in the Fanconi Anemia Cancer Susceptibility Syndrome Leads to Spontaneous Pro-inflammatory Cytokine Production. EBioMedicine *8*, 184–194.

Brewer, M.A. et al. (2005). Exploratory study of ovarian intraepithelial neoplasia. Cancer Epidemiol. Biomarkers Prev. *14*, 299–305.

Britten, R.J., and Davidson, E.H. (1969). Gene Regulation for Higher Cells: A Theory. Science (80-.). *165*, 349–357.

Brocks, D. et al. (2017). DNMT and HDAC inhibitors induce cryptic transcription start sites encoded in long terminal repeats. Nat. Genet. *49*, 1052–1060.

Brouha, B. et al. (2003). Hot L1s account for the bulk of retrotransposition in the human population. Proc. Natl. Acad. Sci. *100*, 5280–5285.

Bürkle, A. (2005). Poly(ADP-ribose): The most elaborate metabolite of NAD+. FEBS J. *272*, 4576–4589.

Burns, K.H. (2017). Transposable elements in cancer. Nat. Rev. Cancer *17*, 415–424.

Burté, F. et al. (2015). Disturbed mitochondrial dynamics and neurodegenerative disorders. Nat. Rev. Neurol. *11*, 11–24.

Cantó, C. et al. (2015). NAD+ Metabolism and the Control of Energy Homeostasis: A Balancing Act between Mitochondria and the Nucleus. Cell Metab. *22*, 31–53.

Cao, H. et al. (2016). IL-13/STAT6 signaling plays a critical role in the epithelialmesenchymal transition of colorectal cancer cells. Oncotarget *7*, 61183–61198.

Carbon, S. et al. (2019). The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res. *47*, D330–D338.

Carlini, F. et al. (2010). The Reverse Transcription Inhibitor Abacavir Shows Anticancer

Activity in Prostate Cancer Cell Lines. PLoS One *5*, e14221.

Carreira, P.E. et al. (2014). L1 retrotransposons, cancer stem cells and oncogenesis. FEBS J. *281*, 63–73.

Carter, V. et al. (2020). High Prevalence and Disease Correlation of Autoantibodies Against p40 Encoded by Long Interspersed Nuclear Elements in Systemic Lupus Erythematosus. Arthritis Rheumatol. *72*, 89–99.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642–655.

Carvill, G.L. et al. (2013). Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. Nat. Genet. *45*, 825–830.

De Cecco, M. et al. (2019). L1 drives IFN in senescent cells and promotes age-associated inflammation. Nature *566*, 73–78.

Chai, Y.L. et al. (1999). The second BRCT domain of BRCA1 proteins interacts with p53 and stimulates transcription from the p21(WAF1/CIP1) promoter. Oncogene *18*, 263–268.

Chang, W. et al. (2005). NuMA is a major acceptor of poly(ADP-ribosyl)ation by tankyrase 1 in mitosis. Biochem. J. *391*, 177–184.

Chen, W. et al. (2009). Direct Interaction between Nrf2 and p21Cip1/WAF1 Upregulates the Nrf2-Mediated Antioxidant Response. Mol. Cell *34*, 663–673.

Chiappinelli, K.B. et al. (2015). Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. Cell *162*, 974–986.

Chiba, N., and Parvin, J.D. (2002). The BRCA1 and BARD1 association with the RNA polymerase II holoenzyme. Cancer Res. *62*, 4222–4228.

Chien, J. et al. (2015). TP53 mutations, tetraploidy and homologous recombination repair

defects in early stage high-grade serous ovarian cancer. Nucleic Acids Res. *43*, 6945–6958.

Choi, J. et al. (2018). Interplay between RNASEH2 and MOV10 controls LINE-1 retrotransposition. Nucleic Acids Res. *46*, 1912–1926.

Chung, N. et al. (2019). Transcriptome analyses of tumor-adjacent somatic tissues reveal genes co-expressed with transposable elements. Mob. DNA *10*, 39.

Chuong, E.B. et al. (2017). Regulatory activities of transposable elements: from conflicts to benefits. Nat. Rev. Genet. *18*, 71–86.

Ciriello, G. et al. (2012). Mutual exclusivity analysis identifies oncogenic network modules. Genome Res. *22*, 398–406.

Cleophas, T.J., and Zwinderman, A.H. (2018). Bayesian Pearson Correlation Analysis. In: Modern Bayesian Statistics in Clinical Research, 111–118.

De Clercq, E. (1993). HIV-1-specific RT inhibitors: Highly selective inhibitors of human immunodeficiency virus type 1 that are specifically targeted at the viral reverse transcriptase. Med. Res. Rev. *13*, 229–258.

Cole, A.J. et al. (2016). Assessing mutant p53 in primary high-grade serous ovarian cancer using immunohistochemistry and massively parallel sequencing. Sci. Rep. *6*, 26191.

Contreras-Galindo, R. et al. (2008). Human Endogenous Retrovirus K (HML-2) Elements in the Plasma of People with Lymphoma and Breast Cancer. J. Virol. *82*, 9329–9336.

Cook, L.S. et al. (2017). Combined oral contraceptive use before the first birth and epithelial ovarian cancer risk. Br. J. Cancer *116*, 265–269.

Cook, P.R., and Tabor, G.T. (2016). Deciphering fact from artifact when using reporter assays to investigate the roles of host factors on L1 retrotransposition. Mob. DNA *7*, 1–15.

Cost, G.J. et al. (2002). Human L1 element target-primed reverse transcription in vitro.

EMBO J. *21*, 5899–5910.

Cost, G.J., and Boeke, J.D. (1998). Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. Biochemistry *37*, 18081–18093.

Coufal, N.G. et al. (2009). L1 retrotransposition in human neural progenitor cells. Nature *460*, 1127–1131.

Coufal, N.G. et al. (2011). Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. Proc. Natl. Acad. Sci. U. S. A. *108*, 20382–20387.

Criscione, S.W. et al. (2014). Transcriptional landscape of repetitive elements in normal and cancer human cells. BMC Genomics *15*, 583.

Crum, C.P. et al. (2007). The distal fallopian tube: A new model for pelvic serous carcinogenesis. Curr. Opin. Obstet. Gynecol. *19*, 3–9.

Cui, Y., and Guo, G. (2016). Immunomodulatory Function of the Tumor Suppressor p53 in Host Immune Response and the Tumor Microenvironment. Int. J. Mol. Sci. *17*, 1942.

Cybulska, P. et al. (2019). Molecular profiling and molecular classification of endometrioid ovarian carcinomas. Gynecol. Oncol. *154*, 516–523.

Dagogo-Jack, I., and Shaw, A.T. (2018). Tumour heterogeneity and resistance to cancer therapies. Nat. Rev. Clin. Oncol. *15*, 81–94.

Deininger, P. (2011). Alu elements: Know the SINEs. Genome Biol. *12*, 236.

Deininger, P., and Belancio, V.P. (2016). Detection of LINE-1 RNAs by Northern Blot. In: Methods in Molecular Biology, 223–236.

Denli, A.M. et al. (2015). Primate-Specific ORF0 Contributes to Retrotransposon-

Mediated Diversity. Cell *163*, 583–593.

Desai, N. et al. (2017). Diverse repetitive element RNA expression defines epigenetic and immunologic features of colon cancer. JCI Insight *2*.

Dewannieux, M. et al. (2003). LINE-mediated retrotransposition of marked Alu sequences. Nat. Genet. *35*, 41–48.

Doherty, J.A. et al. (2017). Challenges and Opportunities in Studying the Epidemiology of Ovarian Cancer Subtypes. Curr. Epidemiol. Reports *4*, 211–220.

Domcke, S. et al. (2013). Evaluating cell lines as tumour models by comparison of genomic profiles. Nat. Commun. *4*, 2126.

DONG, L. et al. (2013). Molecular cloning of Tupaia belangeri chinensis neuropeptide Y and homology comparison with other analogues from primates. Zool. Res. *33*, 75–78.

Doucet-O'Hare, T.T. et al. (2015). LINE-1 expression and retrotransposition in Barrett's esophagus and esophageal carcinoma. Proc. Natl. Acad. Sci. *112*, E4894–E4900.

Doucet-O'Hare, T.T. et al. (2016). Somatically Acquired LINE-1 Insertions in Normal Esophagus Undergo Clonal Expansion in Esophageal Squamous Cell Carcinoma. Hum. Mutat. *37*, 942–954.

Doucet, A.J. et al. (2010). Characterization of LINE-1 ribonucleoprotein particles. PLoS Genet. *6*, 1–19.

Doucet, A.J. et al. (2015). A 3′ Poly(A) Tract Is Required for LINE-1 Retrotransposition. Mol. Cell *60*, 728–741.

Dunham, I. et al. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Dunning, A.M. et al. (1997). Common BRCA1 variants and susceptibility to breast and

ovarian cancer in the general population. Hum. Mol. Genet. *6*, 285–289.

Ecco, G. et al. (2017). KRAB zinc finger proteins. Development *144*, 2719–2729.

Elstrodt, F. et al. (2006). BRCA1 mutation analysis of 41 human breast cancer cell lines reveals three new deleterious mutants. Cancer Res. *66*, 41–45.

Elyanow, R. et al. (2018). Identifying structural variants using linked-read sequencing data. Bioinformatics *34*, 353–360.

Escribano-Díaz, C. et al. (2013). A Cell Cycle-Dependent Regulatory Circuit Composed of 53BP1-RIF1 and BRCA1-CtIP Controls DNA Repair Pathway Choice. Mol. Cell *49*, 872–883.

Esnault, C. (2006). Dual inhibitory effects of APOBEC family proteins on retrotransposition of mammalian endogenous retroviruses. Nucleic Acids Res. *34*, 1522–1531.

Etemadmoghadam, D. et al. (2017). EIF1AX and NRAS mutations co-occur and cooperate in low-grade serous ovarian carcinomas. Cancer Res. *77*, 4268–4278.

Evrony, G.D. et al. (2012). Single-neuron sequencing analysis of l1 retrotransposition and somatic mutation in the human brain. Cell *151*, 483–496.

Evrony, G.D. et al. (2016). Resolving rates of mutation in the brain using single-neuron genomics. Elife *5*.

Ewing, A.D. et al. (2015). Widespread somatic L1 retrotransposition occurs early during gastrointestinal cancer evolution. Genome Res. *25*, 1536–1545.

Ewing, A.D. (2015). Transposable element detection from whole genome sequence data. Mob. DNA *6*, 24.

Fabbro, M. et al. (2004). BARD1 regulates BRCA1 apoptotic function by a mechanism

involving nuclear retention. Exp. Cell Res. *298*, 661–673.

Falandry, C. et al. (2019). EWOC-1: A randomized trial to evaluate the feasibility of three different first-line chemotherapy regimens for vulnerable elderly women with ovarian cancer (OC): A GCIG-ENGOT-GINECO study. J. Clin. Oncol. *37*, 5508–5508.

Fang, E.F. et al. (2016). Nuclear DNA damage signalling to mitochondria in ageing. Nat. Rev. Mol. Cell Biol. *17*, 308–321.

Farmer, H. et al. (2005). Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature *434*, 917–921.

Faulkner, G.J. et al. (2009). The regulated retrotransposon transcriptome of mammalian cells. Nat. Genet. *41*, 563–571.

Fazzari, M.J., and Greally, J.M. (2004). Epigenomics: beyond CpG islands. Nat. Rev. Genet. *5*, 446–455.

Feng, Q. et al. (1996). Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell *87*, 905–916.

Feng, Y. et al. (2017). Deamination-independent restriction of LINE-1 retrotransposition by APOBEC3H. Sci. Rep. *7*, 10881.

Ferlay, J. et al. (2021). Cancer statistics for the year 2020: An overview. Int. J. Cancer ijc.33588.

Feuk, L. et al. (2006). Structural variation in the human genome. Nat. Rev. Genet. *7*, 85–97.

Findlay, G.M. et al. (2018). Accurate classification of BRCA1 variants with saturation genome editing. Nature *562*, 217–222.

Folmes, C.D.L. et al. (2011). Somatic oxidative bioenergetics transitions into pluripotency-

dependent glycolysis to facilitate nuclear reprogramming. Cell Metab. *14*, 264–271.

Fort, M.M. et al. (2001). IL-25 Induces IL-4, IL-5, and IL-13 and Th2-associated pathologies in vivo. Immunity *15*, 985–995.

Foulkes, W.D., and Shuen, A.Y. (2013). In Brief: BRCA1 and BRCA2. J. Pathol. *230*, 347–349.

Frank, O. et al. (2005). Human Endogenous Retrovirus Expression Profiles in Samples from Brains of Patients with Schizophrenia and Bipolar Disorders. J. Virol. *79*, 10890–10901.

Fries, G.R. et al. (2019). Revisiting inflammation in bipolar disorder. Pharmacol. Biochem. Behav. *177*, 12–19.

Fujisawa, T. et al. (2012). IL-13 regulates cancer invasion and metastasis through IL-13Rα2 via ERK/AP-1 pathway in mouse model of human ovarian cancer. Int. J. Cancer *131*, 344–356.

Fung, K.Y. et al. (2013). Interferon-ε protects the female reproductive tract from viral and bacterial infection. Science *339*, 1088–1092.

Furuta, S. et al. (2011). IL-25 Causes Apoptosis of IL-25R-Expressing Breast Cancer Cells Without Toxicity to Nonmalignant Cells. Sci. Transl. Med. *3*, 78ra31-78ra31.

Gaggianesi, M. et al. (2017). IL4 primes the dynamics of breast cancer progression via DUSP4 inhibition. Cancer Res. *77*, 3268–3279.

Galon, J., and Bruni, D. (2019). Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. Nat. Rev. Drug Discov. *18*, 197–218.

Gamwell, L.F. et al. (2012). The Mouse Ovarian Surface Epithelium Contains a Population of LY6A (SCA-1) Expressing Progenitor Cells That Are Regulated by Ovulation-

Associated Factors1. Biol. Reprod. *87*, 80.

Ganesan, S. et al. (2002). BRCA1 supports XIST RNA concentration on the inactive X chromosome. Cell *111*, 393–405.

Ganguly, A. et al. (2003). Exon skipping caused by an intronic insertion of a young Alu Yb9 element leads to severe hemophilia A. Hum. Genet. *113*, 348–352.

Gardini, A. et al. (2014). Genome-wide analysis reveals a role for BRCA1 and PALB2 in transcriptional co-activation. EMBO J. *33*, 890–905.

Gardner, E.J. et al. (2017). The mobile element locator tool (MELT): Population-scale mobile element discovery and biology. Genome Res. *27*, 1916–1929.

Gasior, S.L. et al. (2006). The human LINE-1 retrotransposon creates DNA double-strand breaks. J. Mol. Biol. *357*, 1383–1393.

Goel, S. et al. (2017). CDK4/6 inhibition triggers anti-tumour immunity. Nature *548*, 471–475.

Goerner-Potvin, P., and Bourque, G. (2018). Computational tools to unmask transposable elements. Nat. Rev. Genet. *19*, 688–704.

Göke, J., and Ng, H.H. (2016). CTRL + INSERT : retrotransposons and their contribution to regulation and innovation of the transcriptome. EMBO Rep. *17*, 1131–1144.

Goldstein, I. et al. (2011). Understanding wild-type and mutant p53 activities in human cancer: New landmarks on the way to targeted therapies. Cancer Gene Ther. *18*, 2–11.

Goodier, J.L. et al. (2013). Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. Nucleic Acids Res. *41*, 7401–7419.

Goodier, J.L. (2016). Restricting retrotransposons: A review. Mob. DNA *7*, 16.

Grandi, F.C. et al. (2015). Retrotransposition creates sloping shores: a graded influence of

hypomethylated CpG islands on flanking CpG sites. Genome Res. *25*, 1135–1146.

Grossman, R.L. et al. (2016). Toward a Shared Vision for Cancer Genomic Data. N. Engl. J. Med. *375*, 1109–1112.

Grow, E.J. et al. (2015). Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. Nature *522*, 221–246.

Grundy, E.E. et al. (2021). Transposable element regulation and expression in cancer. FEBS J. febs.15722.

Guler, G.D. et al. (2017). Repression of Stress-Induced LINE-1 Expression Protects Cancer Cell Subpopulations from Lethal Drug Exposure. Cancer Cell *32*, 221-237 e13.

Guo, H. et al. (2014). Autophagy supports genomic stability by degrading retrotransposon RNA. Nat. Commun. *5*, 5276.

Guo, H.L. et al. (2012). The Axin/TNKS complex interacts with KIF3A and is required for insulin-stimulated GLUT4 translocation. Cell Res. *22*, 1246–1257.

Gupte, R. et al. (2017). PARPs and ADP-ribosylation: recent advances linking molecular functions to biological outcomes. Genes Dev. *31*, 101–126.

Hakem, R. et al. (1996). The tumor suppressor gene Brca1 is required for embryonic cellular proliferation in the mouse. Cell *85*, 1009–1023.

Hakes, T.B. et al. (1992). Randomized prospective trial of 5 versus 10 cycles of cyclophosphamide, doxorubicin, and cisplatin in advanced ovarian carcinoma. Gynecol. Oncol. *45*, 284–289.

Hall, J.M. et al. (1990). Linkage of early-onset familial breast cancer to chromosome 17q21. Science (80-.). *250*, 1684–1689.

Hamperl, S., and Cimprich, K.A. (2014). The contribution of co-transcriptional RNA:

DNA hybrid structures to DNA damage and genome instability. DNA Repair (Amst). *19*, 84–94.

Han, J.S. et al. (2004). Transcriptional disruption by the L1 retrotransposon and implications for mammalian transcriptomes. Nature *429*, 268–274.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of Cancer: The Next Generation. Cell *144*, 646–674.

Hancks, D.C., and Kazazian, H.H. (2012). Active human retrotransposons: variation and disease. Curr. Opin. Genet. Dev. *22*, 191–203.

Harris, C.R. et al. (2009). P53 responsive elements in human retrotransposons. Oncogene *28*, 3857–3865.

Hart, T. et al. (2015). High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. Cell *163*, 1515–1526.

Harter, P. et al. (2019). A Randomized Trial of Lymphadenectomy in Patients with Advanced Ovarian Neoplasms. N. Engl. J. Med. *380*, 822–832.

Hashizume, R. et al. (2001). The RING Heterodimer BRCA1-BARD1 Is a Ubiquitin Ligase Inactivated by a Breast Cancer-derived Mutation. J. Biol. Chem. *276*, 14537–14540.

Hastings, P.J. et al. (2009). Mechanisms of change in gene copy number. Nat. Rev. Genet. *10*, 551–564.

Hatchi, E. et al. (2015). BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. Mol. Cell *57*, 636–647.

Helman, E. et al. (2014). Somatic retrotransposition in human cancer revealed by whole-genome and exome sequencing. Genome Res. *24*, 1053–1063.

Henderson, B.R. (2005). Regulation of BRCA1, BRCA2 and BARD1 intracellular

trafficking. BioEssays *27*, 884–893.

Hendrickx, W. et al. (2017). Identification of genetic determinants of breast cancer immune phenotypes by integrative genome-scale analysis. Oncoimmunology *6*, e1253654.

Heras, S.R. et al. (2013). The Microprocessor controls the activity of mammalian retrotransposons. Nat. Struct. Mol. Biol. *20*, 1173–1183.

Herrmann, A. et al. (2018). The SAMHD1-mediated block of LINE-1 retroelements is regulated by phosphorylation. Mob. DNA *9*, 11.

Hess, L.M. et al. (2019). Relationship between progression-free survival and overall survival in randomized clinical trials of targeted and biologic agents in oncology. J. Cancer *10*, 3717–3727.

Hilpert, F. et al. (2007). Feasibility, toxicity and quality of life of first-line chemotherapy with platinum/paclitaxel in elderly patients aged ≥70 years with advanced ovarian cancer - A study by the AGO OVAR Germany. Ann. Oncol. *18*, 282–287.

Horn, A. V. et al. (2014). Human LINE-1 restriction by APOBEC3C is deaminase independent and mediated by an ORF1p interaction that affects LINE reverse transcriptase activity. Nucleic Acids Res. *42*, 396–416.

Houtkooper, R.H. et al. (2013). Mitonuclear protein imbalance as a conserved longevity mechanism. Nature *497*, 451–457.

Huang, H. et al. (2018). IPTMnet: An integrated resource for protein post-translational modification network discovery. Nucleic Acids Res. *46*, D542–D550.

Huang, S.M.A. et al. (2009). Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. Nature *461*, 614–620.

Huen, M.S.Y. et al. (2010). BRCA1 and its toolbox for the maintenance of genome

integrity. Nat. Rev. Mol. Cell Biol. *11*, 138–148.

Huh, J.W. et al. (2009). Gain of new exons and promoters by lineage-specific transposable elements-integration and conservation event on CHRM3 gene. Mol. Cells *28*, 111–117.

Hukezalie, K.R. et al. (2012). In Vitro and Ex Vivo Inhibition of Human Telomerase by Anti-HIV Nucleoside Reverse Transcriptase Inhibitors (NRTIs) but Not by Non-NRTIs. PLoS One *7*, e47505.

Hung, T. et al. (2015). The Ro60 autoantigen binds endogenous retroelements and regulates inflammatory gene expression. Science (80-.). *350*, 455–459.

Hunter, S.M. et al. (2015). Molecular profiling of low grade serous ovarian tumours identifies novel candidate driver genes. Oncotarget *6*, 37663–37667.

Hur, K. et al. (2014). Hypomethylation of long interspersed nuclear element-1 (LINE-1) leads to activation of proto-oncogenes in human colorectal cancer metastasis. Gut *63*, 635–646.

van den Hurk, J.A.J.M. et al. (2007). L1 retrotransposition can occur early in human embryonic development. Hum. Mol. Genet. *16*, 1587–1592.

Ikeda, K. et al. (2015). Identification of estrogen-responsive genes based on the DNA binding properties of estrogen receptors using high-throughput sequencing technology. Acta Pharmacol. Sin. *36*, 24–31.

Imbeault, M. et al. (2017). KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. Nature *543*, 550–554.

Ishak, C.A. et al. (2018). Deregulation of Retroelements as an Emerging Therapeutic Opportunity in Cancer. Trends in Cancer *4*, 583–597.

Iskow, R.C. et al. (2010). Natural Mutagenesis of Human Genomes by Endogenous

Retrotransposons. Cell *141*, 1253–1261.

Jackson, S.P., and Bartek, J. (2009). The DNA-damage response in human biology and disease. Nature *461*, 1071–1078.

Jakymiw, A. et al. (2005). Disruption of GW bodies impairs mammalian RNA interference. Nat. Cell Biol. *7*, 1167–1174.

Jayson, G.C. et al. (2014). Seminar Ovarian cancer. Lancet *384*, 1376–1388.

Jern, P., and Coffin, J.M. (2008). Effects of Retroviruses on Host Genome Function. Annu. Rev. Genet. *42*, 709–732.

Jiang, J. et al. (2011). p53-dependent BRCA1 nuclear export controls cellular susceptibility to DNA damage. Cancer Res. *71*, 5546–5557.

Jiang, Z. et al. (2017). IL-25 blockade inhibits metastasis in breast cancer. Protein Cell *8*, 191–201.

St. John, J.C. et al. (2005). The expression of mitochondrial DNA transcription factors during early cardiomyocyte in vitro differentiation from human embryonic stemn cells. Cloning Stem Cells *7*, 141–153.

Johnston, R. et al. (2016). The identification of a novel role for BRCA1 in regulating RNA polymerase I transcription. Oncotarget *7*, 68097–68110.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. *13*, 484–492.

Jones, R.B. et al. (2008). Nucleoside Analogue Reverse Transcriptase Inhibitors Differentially Inhibit Human LINE-1 Retrotransposition. PLoS One *3*, e1547.

Jones, S. et al. (2012). Low-grade serous carcinomas of the ovary contain very few point mutations. J. Pathol. *226*, 413–420.

Jung, H. et al. (2018). Immune signatures correlate with L1 retrotransposition in gastrointestinal cancers. Genome Res. *28*, 1136–1146.

Jurka, J. et al. (2005). Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. *110*, 462–467.

Kaniecki, K. et al. (2018). A change of view: Homologous recombination at single-molecule resolution. Nat. Rev. Genet. *19*, 191–207.

Karras, G.I. et al. (2005). The macro domain is an ADP-ribose binding module. EMBO J. *24*, 1911–1920.

Katsyuba, E. et al. (2020). NAD+ homeostasis in health and disease. Nat. Metab. *2*, 9–31.

Kattah, N.H. et al. (2010). The U1-snRNP complex: Structural properties relating to autoimmune pathogenesis in rheumatic diseases. Immunol. Rev. *233*, 126–145.

Kawamura, Y. et al. (2019). Extracellular vesicles mediate the horizontal transfer of an active LINE-1 retrotransposon. J. Extracell. Vesicles *8*, 1643214.

Kazazian, H.H. et al. (1988). Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature *332*, 164–166.

Kazazian Jr., H.H., and Moran, J. V (2017). Mobile DNA in Health and Disease. N Engl J Med *377*, 361–370.

Kehoe, S. et al. (2015). Primary chemotherapy versus primary surgery for newly diagnosed advanced ovarian cancer (CHORUS): An open-label, randomised, controlled, non-inferiority trial. Lancet *386*, 249–257.

Kerns, J.A. et al. (2008). Positive selection and increased antiviral activity associated with the PARP-containing isoform of human zinc-finger antiviral protein. PLoS Genet. *4*, 0150–0158.

Khacho, M., and Slack, R.S. (2017). Mitochondrial activity in the regulation of stem cell self-renewal and differentiation. Curr. Opin. Cell Biol. *49*, 1–8.

Kim, D.S. et al. (2020). PARPs and ADP-ribosylation in RNA biology: From RNA expression and processing to protein translation and proteostasis. Genes Dev. *34*, 302–320.

Kioi, M. et al. (2005). Expression and targeting of interleukin-4 receptor for primary and advanced ovarian cancer therapy. Cancer Res. *65*, 8388–8396.

Kleiman, F.E. et al. (2005). BRCA1/BARD1 inhibition of mRNA 3′ processing involves targeted degradation of RNA polymerase II. Genes Dev. *19*, 1227–1237.

Koboldt, D.C. et al. (2012). Comprehensive molecular portraits of human breast tumours. Nature *490*, 61–70.

Komarnitsky, P. et al. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. Genes Dev. *14*, 2452–2460.

Kong, Y. et al. (2019). Transposable element expression in tumors is associated with immune infiltration and increased antigenicity. Nat. Commun. *10*, 5228.

Kopera, H.C. et al. (2011). Similarities between long interspersed element-1 (LINE-1) reverse transcriptase and telomerase. Proc. Natl. Acad. Sci. U. S. A. *108*, 20345–20350.

Kroutter, E.N. et al. (2009). The RNA polymerase dictates ORF1 requirement and timing of LINE and SINE retrotransposition. PLoS Genet. *5*, e1000458.

Krug, L. et al. (2017). Retrotransposon activation contributes to neurodegeneration in a Drosophila TDP-43 model of ALS. PLoS Genet. *13*, e1006635.

Krum, S.A. et al. (2003). BRCA1 Associates with Processive RNA Polymerase II. J. Biol. Chem. *278*, 52012–52020.

Kuchenbaecker, K.B. et al. (2017). Risks of breast, ovarian, and contralateral breast cancer

for BRCA1 and BRCA2 mutation carriers. JAMA - J. Am. Med. Assoc. *317*, 2402–2416.

Kulpa, D.A., and Moran, J. V. (2006). Cis-preferential LINE-1 reverse transcriptase activity in ribonucleoprotein particles. Nat. Struct. Mol. Biol. *13*, 655–660.

Kumar, P.P. et al. (2007). SATB1-binding sequences and Alu-like motifs define a unique chromatin context in the vicinity of human immunodeficiency virus type 1 integration sites. J Virol *81*, 5617–5627.

Kunarso, G. et al. (2010). Transposable elements have rewired the core regulatory network of human embryonic stem cells. Nat. Genet. *42*, 631–634.

Kuo, K.T. et al. (2009). Frequent activating mutations of PIK3CA in ovarian clear cell carcinoma. Am. J. Pathol. *174*, 1597–1601.

Kurnit, K.C. et al. (2021). Updates and New Options in Advanced Epithelial Ovarian Cancer Treatment. Obstet. Gynecol. *137*, 108–121.

Kuusisto, K.M. et al. (2013). Copy Number Variation Analysis in Familial BRCA1/2-Negative Finnish Breast and Ovarian Cancer. PLoS One *8*, e71802.

Kwak, H., and Lis, J.T. (2013). Control of transcriptional elongation. Annu. Rev. Genet. *47*, 483–508.

Lander, E.S. et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Landriscina, M. et al. (2007). Anti-Tumor Activity of Non-Nucleosidic Reverse Transcriptase Inhibitors. Curr. Pharm. Des. *13*, 737–747.

Lawrie, T.A. et al. (2015). Adjuvant (post-surgery) chemotherapy for early stage epithelial ovarian cancer. Cochrane Database Syst. Rev. *2015*.

Lee, E. et al. (2012). Landscape of Somatic Retrotransposition in Human Cancers. Science

(80-.). *337*, 967–971.

Lee, Y. et al. (2007). A candidate precursor to serous carcinoma that originates in the distal fallopian tube. J. Pathol. *211*, 26–35.

Leonova, K.I. et al. (2013). p53 cooperates with DNA methylation and a suicidal interferon response to maintain epigenetic silencing of repeats and noncoding RNAs. Proc. Natl. Acad. Sci. *110*, E89–E98.

Leucci, E. et al. (2016). Melanoma addiction to the long non-coding RNA SAMMSON. Nature *531*, 518–522.

Lev-Maor, G. (2003). The Birth of an Alternatively Spliced Exon: 3' Splice-Site Selection in Alu Exons. Science (80-.). *300*, 1288–1291.

Lever, J. et al. (2019). CancerMine: a literature-mined resource for drivers, oncogenes and tumor suppressors in cancer. Nat. Methods *16*, 505–507.

Lheureux, S. et al. (2019). Epithelial ovarian cancer. Lancet.

Li, M., and Yu, X. (2013). Function of BRCA1 in the DNA Damage Response Is Mediated by ADP-Ribosylation. Cancer Cell *23*, 693–704.

Li, P. et al. (2017). Aicardi–Goutières syndrome protein TREX1 suppresses L1 and maintains genome integrity through exonuclease-independent ORF1p depletion. Nucleic Acids Res. *45*, 4619–4631.

Li, S. et al. (2018). Hypomethylation of LINE-1 elements in schizophrenia and bipolar disorder. J. Psychiatr. Res. *107*, 68–72.

Liang, W. et al. (2016). APOBEC3DE Inhibits LINE-1 Retrotransposition by Interacting with ORF1p and Influencing LINE Reverse Transcriptase Activity. PLoS One *11*, e0157220.

Linder, P., and Jankowsky, E. (2011). From unwinding to clamping ĝ€" the DEAD box RNA helicase family. Nat. Rev. Mol. Cell Biol. *12*, 505–516.

Lister, R. et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315–322.

Liu, E.Y. et al. (2019). Loss of Nuclear TDP-43 Is Associated with Decondensation of LINE Retrotransposons. Cell Rep. *27*, 1409-1421.e6.

Liu, J. et al. (2005). A role for the P-body component GW182 in microRNA function. Nat. Cell Biol. *7*, 1161–1166.

Liu, N. et al. (2018). Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. Nature *553*, 228–232.

Llave, C. et al. (2002). Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. Science (80-.). *297*, 2053–2056.

López De Silanes, I. et al. (2004). Identification of a target RNA motif for RNA-binding protein HuR. Proc. Natl. Acad. Sci. U. S. A. *101*, 2987–2992.

Lu, Z., and Chen, J. (2014). Introduction of WHO classification of tumours of female reproductive organs, fourth edition. Chinese J. Pathol. *43*, 649–650.

Luan, D.D. et al. (1993). Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: A mechanism for non-LTR retrotransposition. Cell *72*, 595–605.

Luijsterburg, M.S. et al. (2016). PARP1 Links CHD2-Mediated Chromatin Expansion and H3.3 Deposition to DNA Repair by Non-homologous End-Joining. Mol. Cell *61*, 547–562.

Luo, M.J. et al. (2001). Pre-mrna splicing and mRNA export linked by direct interactions between UAP56 and aly. Nature *413*, 644–647.

MacLennan, M. et al. (2017). Mobilization of LINE-1 retrotransposons is restricted by

Tex19.1 in mouse embryonic stem cells. Elife *6*.

Maddox, J. et al. (2012). Transcription Factor Oct1 Is a Somatic and Cancer Stem Cell Determinant. PLoS Genet. *8*, e1003048.

Malik, H.S. (2000). Poised for Contagion: Evolutionary Origins of the Infectious Abilities of Invertebrate Retroviruses. Genome Res. *10*, 1307–1318.

Mallery, D.L. et al. (2002). Activation of the E3 ligase function of the BRCA1/BARD1 complex by polyubiquitin chains. EMBO J. *21*, 6755–6762.

Malone, C.D. et al. (2014). The exon junction complex controls transposable element activity by ensuring faithful splicing of the piwi transcript. Genes Dev. *28*, 1786–1799.

Malone, C.D., and Hannon, G.J. (2009). Small RNAs as Guardians of the Genome. Cell *136*, 656–668.

Mangiacasale, R. et al. (2003). Exposure of normal and transformed cells to nevirapine, a reverse transcriptase inhibitor, reduces cell growth and promotes differentiation. Oncogene *22*, 2750–2761.

Mark, W.Y. et al. (2005). Characterization of segments from the central region of BRCA1: An intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions? J. Mol. Biol. *345*, 275–287.

Martin, S.L. et al. (2003). Trimeric structure for an essential protein in L1 retrotransposition. Proc. Natl. Acad. Sci. U. S. A. *100*, 13815–13820.

Mason, C.E. et al. (2010). Location analysis for the estrogen receptor-α reveals binding to diverse ERE sequences and widespread binding within repetitive DNA elements. Nucleic Acids Res. *38*, 2355–2368.

Mathias, S. et al. (1991). Reverse transcriptase encoded by a human transposable element.

Science (80-.). *254*, 1808–1810.

Matoba, S. et al. (2006). p53 regulates mitochondrial respiration. Science (80-.). *312*, 1650–1653.

Matsuo, K. et al. (2018). Trends of low-grade serous ovarian carcinoma in the United States. J. Gynecol. Oncol. *29*.

McClintock, B. (1950). The origin and behavior of mutable loci in maize. Proc. Natl. Acad. Sci. *36*, 344–355.

McGuire, W.P. et al. (1996). Cyclophosphamide and Cisplatin Compared with Paclitaxel and Cisplatin in Patients with Stage III and Stage IV Ovarian Cancer. N. Engl. J. Med. *334*, 1–6.

McMillan, J.P., and Singer, M.F. (1993). Translation of the human LINE-1 element, L1Hs. Proc. Natl. Acad. Sci. U. S. A. *90*, 11533–11537.

Meinhold-Heerlein, I. et al. (2015). Statement by the Kommission Ovar of the AGO: The New FIGO and WHO Classifications of Ovarian, Fallopian Tube and Primary Peritoneal Cancer. Geburtshilfe Frauenheilkd. *75*, 1021–1027.

Meinhold-Heerlein, I. et al. (2016). The new WHO classification of ovarian, fallopian tube, and primary peritoneal cancer and its clinical implications. Arch. Gynecol. Obstet. *293*, 695–700.

Menendez, L. et al. (2004). L1 and HERV-W retrotransposons are hypomethylated in human ovarian carcinomas. Mol. Cancer *3*, 12.

Van Meter, M. et al. (2014). SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. Nat. Commun. *5*, 5011.

Miki, Y. et al. (1992). Disruption of the APC gene by a retrotransposal insertion of L1

sequence in a colon cancer. Cancer Res. *52*, 643–645.

Milewska, A. et al. (2018). APOBEC3-mediated restriction of RNA virus replication. Sci. Rep. *8*, 5960.

Mir, A.A. et al. (2015). euL1db: The European database of L1HS retrotransposon insertions in humans. Nucleic Acids Res. *43*, D43–D47.

Mita, P. et al. (2020). BRCA1 and S phase DNA repair pathways restrict LINE-1 retrotransposition in human cells. Nat. Struct. Mol. Biol. *27*, 179–191.

Mita, P., and Boeke, J.D. (2016). How retrotransposons shape genome regulation. Curr. Opin. Genet. Dev. *37*, 90–100.

Mitsuishi, Y. et al. (2012). The Keap1–Nrf2 system in cancers: stress response and anabolic metabolism. Front. Oncol. *2*.

Miyoshi, T. et al. (2019). Poly(ADP-Ribose) Polymerase 2 Recruits Replication Protein A to Sites of LINE-1 Integration to Facilitate Retrotransposition. Mol. Cell *75*, 1286-1298.e12.

Moisan, A. et al. (2004). BRCA1 Can Modulate RNA Polymerase II Carboxy-Terminal Domain Phosphorylation Levels. Mol. Cell. Biol. *24*, 6947–6956.

Moldovan, J.B., and Moran, J. V. (2015). The Zinc-Finger Antiviral Protein ZAP Inhibits LINE and Alu Retrotransposition. PLOS Genet. *11*, e1005121.

Montano, M.M. et al. (1998). Transcriptional regulation of the human quinone reductase gene by antiestrogen-liganded estrogen receptor-α and estrogen receptor-β. J. Biol. Chem. *273*, 25443–25449.

Moore, K. et al. (2018). Maintenance Olaparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. N. Engl. J. Med. *379*, 2495–2505.

Moran, J. V et al. (1996). High frequency retrotransposition in cultured mammalian cells. Cell *87*, 917–927.

Mørch, L.S. et al. (2017). Contemporary Hormonal Contraception and the Risk of Breast Cancer. N. Engl. J. Med. *377*, 2228–2239.

Morris, G. et al. (2017). A model of the mitochondrial basis of bipolar disorder. Neurosci. Biobehav. Rev. *74*, 1–20.

Morrish, T.A. et al. (2002). DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nat. Genet. *31*, 159–165.

Morse, B. et al. (1988). Insertional mutagenesis of the myc locus by a LINE-1 sequence in a human breast carcinoma. Nature *333*, 87–90.

Moullan, N. et al. (2015). Tetracyclines Disturb Mitochondrial Function across Eukaryotic Models: A Call for Caution in Biomedical Research. Cell Rep. *10*, 1681–1691.

Moynahan, M.E. et al. (1999). Brca1 controls homology-directed DNA repair. Mol. Cell *4*, 511–518.

Muñoz-Fontela, C. et al. (2008). Transcriptional role of p53 in interferon-mediated antiviral immunity. J. Exp. Med. *205*, 1929–1938.

Muotri, A.R. et al. (2005). Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature *435*, 903–910.

Murachelli, A.G. et al. (2012). The structure of the ASAP core complex reveals the existence of a Pinin-containing PSAP complex. Nat. Struct. Mol. Biol. *19*, 378–386.

Mustelin, T., and Ukadike, K.C. (2020). How Retroviruses and Retrotransposons in Our Genome May Contribute to Autoimmunity in Rheumatological Conditions. Front. Immunol. *11*.

Nadeau, G. et al. (2000). BRCA1 can stimulate gene transcription by a unique mechanism. EMBO Rep. *1*, 260–265.

Nair, S.J. et al. (2016). Genetic suppression reveals DNA repair-independent antagonism between BRCA1 and COBRA1 in mammary gland development. Nat. Commun. *7*, 10913.

Natoli, M. et al. (2021). Transcriptional analysis of multiple ovarian cancer cohorts reveals prognostic and immunomodulatory consequences of ERV expression. J. Immunother. Cancer *9*, e001519.

Nguyen, T.H.M. et al. (2018). L1 Retrotransposon Heterogeneity in Ovarian Tumor Cell Evolution. Cell Rep. *23*, 3730–3740.

Di Nicola, V. (2019). Omentum a powerful biological source in regenerative surgery. Regen. Ther. *11*, 182–191.

Nieto, L. et al. (2007). Differential Effects of Phosphorylation on DNA Binding Properties of N Oct-3 Are Dictated by Protein/DNA Complex Structures. J. Mol. Biol. *370*, 687–700.

Obinata, D. et al. (2012). Oct1 regulates cell growth of LNCaP cells and is a prognostic factor for prostate cancer. Int. J. Cancer *130*, 1021–1028.

Onda, T. et al. (2020). Comparison of survival between primary debulking surgery and neoadjuvant chemotherapy for stage III/IV ovarian, tubal and peritoneal cancers in phase III randomised trial. Eur. J. Cancer *130*, 114–125.

Onozawa, M. et al. (2014). Repair of DNA double-strand breaks by templated nucleotide sequence insertions derived from distant regions of the genome. Proc. Natl. Acad. Sci. *111*, 7729–7734.

Ostertag, E.M. et al. (2000). Determination of L1 retrotransposition kinetics in cultured cells. Nucleic Acids Res. *28*, 1418–1423.

Ouchi, T. et al. (2000). Collaboration of signal transducer and activator of transcription 1 (STAT1) and BRCA1 in differential regulation of IFN-γ target genes. Proc. Natl. Acad. Sci. U. S. A. *97*, 5208–5213.

Ozols, R.F. et al. (2003). Phase III Trial of Carboplatin and Paclitaxel Compared With Cisplatin and Paclitaxel in Patients With Optimally Resected Stage III Ovarian Cancer: A Gynecologic Oncology Group Study. J. Clin. Oncol. *21*, 3194–3200.

Pace, J.K., and Feschotte, C. (2007). The evolutionary history of human DNA transposons: Evidence for intense activity in the primate lineage. Genome Res. *17*, 422–432.

Parasar, P. et al. (2017). Endometriosis: Epidemiology, Diagnosis and Clinical Management. Curr. Obstet. Gynecol. Rep. *6*, 34–41.

Park, K.S. et al. (2002). FCCP depolarizes plasma membrane potential by activating proton and Na+ currents in bovine aortic endothelial cells. Pflugers Arch. Eur. J. Physiol. *443*, 344–352.

Parmar, M.K.B. et al. (1998). ICON2: Randomised trial of single-agent carboplatin against three-drug combination of CAP (cyclophosphamide, doxorubicin, and cisplatin) in women with ovarian cancer. Lancet *352*, 1571–1576.

Pauklin, S. et al. (2009). Estrogen directly activates AID transcription and function. J. Exp. Med. *206*, 99–111.

Penzkofer, T. et al. (2017). L1Base 2: More retrotransposition-active LINE-1s, more mammalian genomes. Nucleic Acids Res. *45*, D68–D73.

Pépin, D. et al. (2015). AAV9 delivering a modified human Mullerian inhibiting substance as a gene therapy in patient-derived xenografts of ovarian cancer. Proc. Natl. Acad. Sci. U. S. A. *112*, E4418–E4427.

Pereira, G.C. et al. (2018). Properties of LINE-1 proteins and repeat element expression in the context of amyotrophic lateral sclerosis. Mob. DNA *9*, 35.

Pisanic, T.R. et al. (2019). Long Interspersed Nuclear Element 1 Retrotransposons Become Deregulated during the Development of Ovarian Cancer Precursor Lesions. Am. J. Pathol. *189*, 513–520.

Pizarro, J.G., and Cristofari, G. (2016). Post-Transcriptional Control of LINE-1 Retrotransposition by Cellular Host Factors in Somatic Cells. Front. Cell Dev. Biol. *4*.

Prakash, R. et al. (2015). Homologous Recombination and Human Health: The Roles of BRCA1, BRCA2, and Associated Proteins. Cold Spring Harb. Perspect. Biol. *7*, a016600.

Raspagliesi, F. et al. (2013). Advanced ovarian cancer: Omental bursa, lesser omentum, celiac, portal and triad nodes spread as cause of inaccurate evaluation of residual tumor. Gynecol. Oncol. *129*, 92–96.

Raudvere, U. et al. (2019). G:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. *47*, W191–W198.

Rauh-Hain, J.A. et al. (2017). Overall survival following neoadjuvant chemotherapy vs primary cytoreductive surgery in women with epithelial ovarian cancer: Analysis of the National Cancer Database. JAMA Oncol. *3*, 76–82.

Ravindran, S. (2012). Barbara McClintock and the discovery of jumping genes. Proc. Natl. Acad. Sci. *109*, 20198–20199.

Reid, B.M. et al. (2017). Epidemiology of ovarian cancer: a review. Cancer Biol. Med. *14*, 9–32.

Reuss, A. et al. (2019). TRUST: Trial of Radical Upfront Surgical Therapy in advanced ovarian cancer (ENGOT ov33/AGO-OVAR OP7). Int. J. Gynecol. Cancer *29*, 1327–1331.

Richardson, S.R. et al. (2014). APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition. Elife *3*.

Ricotta, G. et al. (2021). Brenner Borderline Ovarian Tumor: A Case Series and Literature Review. Ann. Surg. Oncol.

Rio, D.C. (2015). Northern blots: Capillary transfer of RNA from Agarose gels and filter hybridization using standard stringency conditions. Cold Spring Harb. Protoc. *2015*, 306–313.

Rishishwar, L. et al. (2017). Benchmarking computational tools for polymorphic transposable element detection. Brief. Bioinform. *18*, 908–918.

Rodić, N. et al. (2014). Long Interspersed Element-1 Protein Expression Is a Hallmark of Many Human Cancers. Am. J. Pathol. *184*, 1280–1286.

Rodić, N. et al. (2015). Retrotransposon insertions in the clonal evolution of pancreatic ductal adenocarcinoma. Nat. Med. *21*, 1060–1064.

Rodić, N., and Burns, K.H. (2013). Long Interspersed Element–1 (LINE-1): Passenger or Driver in Human Neoplasms? PLoS Genet. *9*, e1003402.

Rodova, M. et al. (2013). CMV promoter is repressed by p53 and activated by JNK pathway. Plasmid *69*, 223–230.

Rodriguez-Martin, B. et al. (2020). Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition. Nat. Genet. *52*, 306–319.

Roers, A. et al. (2016). Recognition of Endogenous Nucleic Acids by the Innate Immune System. Immunity *44*, 739–754.

Rooney, M.S. et al. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell *160*, 48–61.

Roulois, D. et al. (2015). DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. Cell *162*, 961–973.

Rudaitis, V. et al. (2014). BRCA1/2 mutation status is an independent factor of improved survival for advanced (stage III-IV) ovarian cancer. Int. J. Gynecol. Cancer *24*, 1395–1400.

Rudin, C.M., and Thompson, C.B. (2001). Transcriptional activation of short interspersed elements by DNA-damaging agents. Genes Chromosom. Cancer *30*, 64–71.

Ruffner, H. et al. (2001). Cancer-predisposing mutations within the RING domain of BRCA1: Loss of ubiquitin protein ligase activity and protection from radiation hypersensitivity. Proc. Natl. Acad. Sci. U. S. A. *98*, 5134–5139.

Ruffner, H., and Verma, I.M. (1997). BRCA1 is a cell cycle-regulated nuclear phosphoprotein. Proc. Natl. Acad. Sci. U. S. A. *94*, 7138–7143.

Ryu, K.W. et al. (2015). New facets in the regulation of gene expression by ADP-ribosylation and poly(ADP-ribose) polymerases. Chem. Rev. *115*, 2453–2481.

Saito, Y. et al. (2016). Inhibition of DNA Methylation Suppresses Intestinal Tumor Organoids by Inducing an Anti-Viral Response. Sci. Rep. *6*, 25311.

Sánchez-Taltavull, D. et al. (2016). Bayesian Correlation Analysis for Sequence Count Data. PLoS One *11*, e0163595.

Santivasi, W.L. et al. (2015). Association between cytosolic expression of BRCA1 and metastatic risk in breast cancer. Br. J. Cancer *113*, 453–459.

Sanz, L.A., and Chédin, F. (2019). High-resolution, strand-specific R-loop mapping via S9.6-based DNA–RNA immunoprecipitation and high-throughput sequencing. Nat. Protoc. *14*, 1734–1755.

Sauta, E. et al. (2020). A Bayesian data fusion based approach for learning genome-wide

transcriptional regulatory networks. BMC Bioinformatics *21*.

Sayaman, R.W. et al. (2021). Germline genetic contribution to the immune landscape of cancer. Immunity *54*, 367-386.e8.

Schauer, S.N. et al. (2018). L1 retrotransposition is a common feature of mammalian hepatocarcinogenesis. Genome Res. *28*, 639–653.

Schuster, H. et al. (2017). The immunopeptidomic landscape of ovarian carcinomas. Proc Natl Acad Sci U S A *114*, E9942–E9951.

Sciamanna, I. et al. (2005). Inhibition of endogenous reverse transcriptase antagonizes human tumor growth. Oncogene *24*, 3923–3931.

Sciamanna, I. et al. (2016). The Reverse Transcriptase Encoded by LINE-1 Retrotransposons in the Genesis, Progression, and Therapy of Cancer. Front. Chem. *4*.

Scott, A.F. et al. (1987). Origin of the human L1 elements: Proposed progenitor genes deduced from a consensus DNA sequence. Genomics *1*, 113–125.

Scott, E., and Devine, S. (2017). The Role of Somatic L1 Retrotransposition in Human Cancers. Viruses *9*, 131.

Scott, E.C. et al. (2016). A hot L1 retrotransposon evades somatic repression and initiates human colorectal cancer. Genome Res. *26*, 745–755.

Scully, R. et al. (1997a). Association of BRCA1 with Rad51 in mitotic and meiotic cells. Cell *88*, 265–275.

Scully, R. et al. (1997b). BRCA1 is a component of the RNA polymerase II holoenzyme. Proc. Natl. Acad. Sci. U. S. A. *94*, 5605–5610.

Sela, N. et al. (2007). Comparative analysis of transposed element insertion within human and mouse genomes reveals Alu's unique role in shaping the human transcriptome.

Genome Biol. *8*, R127.

Sen, S.K. et al. (2007). Endonuclease-independent insertion provides an alternative pathway for L1 retrotransposition in the human genome. Nucleic Acids Res. *35*, 3741–3751.

Servant, G. et al. (2017). The Nucleotide Excision Repair Pathway Limits L1 Retrotransposition. Genetics *205*, 139–153.

Seth, R.B. et al. (2005). Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-κB and IRF3. Cell *122*, 669–682.

Setrerrahmane, S., and Xu, H. (2017). Tumor-related interleukins: Old validated targets for new anti-cancer drug development. Mol. Cancer *16*, 153.

Sharma, S. et al. (2010). Epigenetics in cancer. Carcinogenesis *31*, 27–36.

Shen, H. et al. (2008). Distinct activities of the DExD/H-box splicing factor hUAP56 facilitate stepwise assembly of the spliceosome. Genes Dev. *22*, 1796–1803.

Shen, Y.J. et al. (2015). Genome-derived cytosolic DNA mediates type I interferon-dependent rejection of B cell lymphoma cells. Cell Rep. *11*, 460–473.

Shih-Chu Ho, E. et al. (2001). p53 mutation is infrequent in clear cell carcinoma of the ovary. Gynecol. Oncol. *80*, 189–193.

Shukla, R. et al. (2013a). Endogenous retrotransposition activates oncogenic pathways in hepatocellular carcinoma. Cell *153*, 101–111.

Shukla, R. et al. (2013b). Endogenous Retrotransposition Activates Oncogenic Pathways in Hepatocellular Carcinoma. Cell *153*, 101–111.

Silva, T.C. et al. (2016). TCGA Workflow: Analyze cancer genomics and epigenomics data using Bioconductor packages. F1000Research *5*, 1542.

Sinibaldi-Vallebona, P. et al. (2011). Retrotransposon-Encoded Reverse Transcriptase in the Genesis, Progression and Cellular Plasticity of Human Cancer. Cancers (Basel). *3*, 1141–1157.

Siomi, M.C. et al. (2011). PIWI-interacting small RNAs: the vanguard of genome defence. Nat. Rev. Mol. Cell Biol. *12*, 246–258.

Skourti-Stathaki, K., and Proudfoot, N.J. (2014). A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes Dev. *28*, 1384–1396.

Slomovitz, B. et al. (2020). Low-grade serous ovarian cancer: State of the science. Gynecol. Oncol. *156*, 715–725.

Smith, S., and De Lange, T. (2000). Tankyrase promotes telomere elongation in human cells. Curr. Biol. *10*, 1299–1302.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nat. Rev. Genet. *14*, 204–220.

So, B.R. et al. (2016). A U1 snRNP-specific assembly pathway reveals the SMN complex as a versatile hub for RNP exchange. Nat. Struct. Mol. Biol. *23*, 225–230.

Solyom, S. et al. (2012). Extensive somatic L1 retrotransposition in colorectal tumors. Genome Res. *22*, 2328–2338.

Stertz, L. et al. (2015). Damage-associated molecular patterns and immune activation in bipolar disorder. Acta Psychiatr. Scand. *132*, 211–217.

Stetson, D.B. et al. (2008). Trex1 Prevents Cell-Intrinsic Initiation of Autoimmunity. Cell.

Stordal, B. et al. (2013). BRCA1/2 mutation analysis in 41 ovarian cell lines reveals only one functionally deleterious BRCA1 mutation. Mol. Oncol. *7*, 567–579.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41–45.

Sudmant, P.H. et al. (2015). An integrated map of structural variation in 2,504 human genomes. Nature *526*, 75–81.

Sultana, T. et al. (2019). The Landscape of L1 Retrotransposons in the Human Genome Is Shaped by Pre-insertion Sequence Biases and Post-insertion Selection. Mol. Cell *74*, 555-570.e7.

Sun, J. et al. (2011). Genetic and genomic analyses of RNA polymerase II-pausing factor in regulation of mammalian transcription and cell growth. J. Biol. Chem. *286*, 36248–36257.

Sun, X. et al. (2018). Transcription factor profiling reveals molecular choreography and key regulators of human retrotransposon expression. Proc. Natl. Acad. Sci. U. S. A. *115*, E5526–E5535.

Suzuki, J. et al. (2009). Genetic evidence that the non-homologous end-joining repair pathway is involved in LINE retrotransposition. PLoS Genet. *5*, e1000461.

Swanton, C. et al. (2015). APOBEC Enzymes: Mutagenic Fuel for Cancer Evolution and Heterogeneity. Cancer Discov. *5*, 704–712.

Szak, S.T. et al. (2002). Molecular archeology of L1 insertions in the human genome. Genome Biol. *3*.

Takaoka, M., and Miki, Y. (2018). BRCA1 gene: function and deficiency. Int. J. Clin. Oncol. *23*, 36–44.

Tam, O.H. et al. (2019). Diseases of the nERVous system: retrotransposon activity in neurodegenerative disease. Mob. DNA *10*, 32.

Tang, Z. et al. (2017). Human transposon insertion profiling: Analysis, visualization and identification of somatic LINE-1 insertions in ovarian cancer. Proc. Natl. Acad. Sci. U. S. A. *114*, E733–E740.

Tani, H. et al. (2012). Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. Genome Res *22*, 947–956.

Taylor, M.S. et al. (2013). Affinity Proteomics Reveals Human Host Factors Implicated in Discrete Stages of LINE-1 Retrotransposition. Cell *155*, 1034–1048.

Tharp, M.E. et al. (2020). Maximizing the ovarian reserve in mice by evading LINE-1 genotoxicity. Nat. Commun. *11*, 330.

Thornburg, B.G. et al. (2006). Transposable elements as a significant source of transcription regulating signals. In: Gene, 104–110.

To, M.S. et al. (2010). Mitochondrial uncoupler FCCP activates proton conductance but does not block store-operated Ca2+ current in liver cells. Arch. Biochem. Biophys. *495*, 152–158.

Treiber, T. et al. (2019). Regulation of microRNA biogenesis and its crosstalk with other cellular pathways. Nat. Rev. Mol. Cell Biol. *20*, 5–20.

Tubbs, A., and Nussenzweig, A. (2017). Endogenous DNA Damage as a Source of Genomic Instability in Cancer. Cell *168*, 644–656.

Tubio, J.M.C. et al. (2014). Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. Science (80-.). *345*, 1251343–1251343.

Tudrej, P. et al. (2018). Establishment and Characterization of the Novel High-Grade Serous Ovarian Cancer Cell Line OVPA8. Int. J. Mol. Sci. *19*, 2080.

Upton, K.R. et al. (2015). Ubiquitous L1 mosaicism in hippocampal neurons. Cell *161*,

228–239.

Vencken, P.M.L.H. et al. (2011). Chemosensitivity and outcome of BRCA1- and BRCA2-associated ovarian cancer patients after first-line chemotherapy compared with sporadic ovarian cancer patients. Ann. Oncol. *22*, 1346–1352.

Veras, I. et al. (2009). Inhibition of RNA Polymerase III Transcription by BRCA1. J. Mol. Biol. *387*, 523–531.

Vergote, I. et al. (2010). Neoadjuvant Chemotherapy or Primary Surgery in Stage IIIC or IV Ovarian Cancer. N. Engl. J. Med. *363*, 943–953.

Viollet, S. et al. (2014). L1 retrotransposition. Mob. Genet. Elements *4*, e28907.

Vyas, S. et al. (2013). A systematic analysis of the PARP protein family identifies new functions critical for cell physiology. Nat. Commun. *4*, 2240.

Wahlberg, E. et al. (2012). Family-wide chemical profiling and structural analysis of PARP and tankyrase inhibitors. Nat. Biotechnol. *30*, 283–288.

Wang, B. et al. (2007). Abraxas and RAP80 form a BRCA1 protein complex required for the DNA damage response. Science (80-.). *316*, 1194–1198.

Wang, J. et al. (2014). Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. Nature *516*, 405–409.

Wang, J. et al. (2020). Retrotransposons in pluripotent stem cells. Cell Regen. *9*.

Wang, P. et al. (2002). Modulation of alternative pre-mRNA splicing in vivo by pinin. Biochem. Biophys. Res. Commun. *294*, 448–455.

Wang, Q. et al. (1998). BRCA1 binds c-Myc and inhibits its transcriptional and transforming activity in cells. Oncogene *17*, 1939–1948.

Wang, X. et al. (2019). Bulk tissue cell type deconvolution with multi-subject single-cell

expression reference. Nat. Commun. *10*, 380.

Wang, Y. et al. (2000). BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. Genes Dev. *14*, 927–939.

Wang, Y. peng et al. (2016). Elevated OCT1 participates in colon tumorigenesis and independently predicts poor prognoses of colorectal cancer patients. Tumor Biol. *37*, 3247–3255.

Weinstein, J.N. et al. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. Nat. Genet. *45*, 1113–1120.

Weischenfeldt, J. et al. (2013). Phenotypic impact of genomic structural variation: insights from and for human disease. Nat. Rev. Genet. *14*, 125–138.

Wiegand, K.C. et al. (2010). ARID1A Mutations in Endometriosis-Associated Ovarian Carcinomas. N. Engl. J. Med. *363*, 1532–1543.

Wiehle, L., and Breiling, A. (2016). Chromatin immunoprecipitation. Methods Mol. Biol. *1480*, 7–21.

Wildschutte, J.H. et al. (2016). Discovery of unfixed endogenous retrovirus insertions in diverse human populations. Proc. Natl. Acad. Sci. *113*, E2326–E2334.

Wilhelm-Benartzi, C.S. et al. (2012). Inutero exposures, infant growth, and DNA methylation of repetitive elements and developmentally related genes in human placenta. Environ. Health Perspect. *120*, 296–302.

Worley, M.J. et al. (2014). What is the optimal treatment for obese patients with advanced ovarian carcinoma? Am. J. Obstet. Gynecol. *211*, 231.e1-231.e9.

Wylie, A. et al. (2016). p53 genes function to restrain mobile elements. Genes Dev. *30*, 64–77.

Xiang, J. et al. (2017). Preoperative monocyte-to-lymphocyte ratio in peripheral blood predicts stages, metastasis, and histological grades in patients with ovarian cancer. Transl. Oncol. *10*, 33–39.

Xiao-Jie, L. et al. (2016). LINE-1 in cancer: multifaceted functions and potential clinical implications. Genet. Med. *18*, 431–439.

Xiao, S. et al. (2014). High expression of octamer transcription factor 1 in cervical cancer. Oncol. Lett. *7*, 1889–1894.

Xiong, Y., and Eickbush, T.H. (1990). Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. *9*, 3353–3362.

Yang, D. et al. (2011). Association of BRCA1 and BRCA2 mutations with survival, chemotherapy sensitivity, and gene mutator phenotype in patients with ovarian cancer. JAMA - J. Am. Med. Assoc. *306*, 1557–1565.

Yang, N. et al. (2003). An important role for RUNX3 in human L1 transcription and retrotransposition. Nucleic Acids Res *31*, 4929–4940.

Young, R.C. et al. (1991). Adjuvant therapy in stage I and stage II epithelial ovarian cancer: Results of two prospective randomized trials. Obstet. Gynecol. Surv. *46*, 111–113.

Yu, F. et al. (2001). Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. Nucleic Acids Res. *29*, 4493–4501.

Yu, P. et al. (2012). Nucleic Acid-Sensing Toll-like Receptors Are Essential for the Control of Endogenous Retrovirus Viremia and ERV-Induced Tumors. Immunity *37*, 867–879.

Zhang, A. et al. (2014). RNase L restricts the mobility of engineered retrotransposons in cultured human cells. Nucleic Acids Res. *42*, 3803–3820.

Zhang, F. et al. (2009). PALB2 Links BRCA1 and BRCA2 in the DNA-Damage Response.

Curr. Biol. *19*, 524–529.

Zhang, H. et al. (1998). BRCA1 physically associates with p53 and stimulates its transcriptional activity. Oncogene *16*, 1713–1721.

Zhang, X. et al. (2017). Attenuation of RNA polymerase II pausing mitigates BRCA1-associated R-loop accumulation and tumorigenesis. Nat. Commun. *8*, 15908.

Zhang, X.H.F., and Chasin, L.A. (2006). Comparison of multiple vertebrate genomes reveals the birth and evolution of human exons. Proc. Natl. Acad. Sci. U. S. A. *103*, 13427–13432.

Zhang, Z. (2016). Variable selection with stepwise and best subset approaches. Ann. Transl. Med. *4*, 136–136.

Zhao, H. et al. (2017). The immunomodulatory anticancer agent, RRx-001, induces an interferon response through epigenetic induction of viral mimicry. Clin. Epigenetics *9*, 4.

Zheng, L. et al. (2001). BRCA1 mediates ligand-independent transcriptional repression of the estrogen receptor. Proc. Natl. Acad. Sci. U. S. A. *98*, 9587–9592.

Zhou, L. et al. (2021). Low-dose carboplatin reprograms tumor immune microenvironment through STING signaling pathway and synergizes with PD-1 inhibitors in lung cancer. Cancer Lett. *500*, 163–171.

Zhu, C. et al. (2013). Structural insight into dGTP-dependent activation of tetrameric SAMHD1 deoxynucleoside triphosphate triphosphohydrolase. Nat. Commun. *4*, 2722.

Zhu, Y. et al. (2011). Zinc-finger antiviral protein inhibits HIV-1 infection by selectively targeting multiply spliced viral mRNAs for degradation. Proc. Natl. Acad. Sci. U. S. A. *108*, 15834–15839.