Circular RNAs in the pathogenesis of cancer

Zhao, Xing

DOI:
10.33612/diss.689854111

Chapter 1

Introduction and scope of the thesis

Outline

1. Circular RNAs
2. Technologies used to study circular RNAs
   2.1 Detection of circular RNAs
   2.2 Gain- and loss-of-function of circular RNAs
   2.3 Identification of molecules binding to circular RNAs
3. B-cell lymphoma
4. Circular RNAs in B-cell lymphoma
5. Breast cancer
6. Circular RNAs in breast cancer
7. Scope of thesis

1. Circular RNAs

Circular (circ)RNAs were initially discovered four decades ago [1] and were considered to be the result of aberrant splicing products [2]. CircRNAs represent a subgroup of the non-coding RNA family with a covalently closed loop structure generated by head-to-tail splicing. Due to the absence of 5' and 3' ends, circRNAs are resistant to degradation by RNases. This results in increased stability of circRNAs as compared to linear RNAs. In addition, circRNAs are enriched in brain, are conserved, and have tissue and differentiation stage-specific expression patterns [3-6]. Various functions have been reported for a small subset of the currently known circRNAs. Cytoplasmic circRNAs serve as miRNA or RNA-binding protein (RBP) decoys, interact with mRNA or proteins, and can be translated into peptides or proteins. Nuclear circRNAs can regulate gene transcription or pre-mRNA splicing. CircRNAs regulate important cellular processes and pathways. Based on aberrant expression patterns and functional studies, circRNAs have also been widely implicated in the pathogenesis of cancer. In **Chapter 2**, a more detailed overview of circRNAs, including biogenesis, regulation of circRNAs by RBPs, other molecular mechanisms, and currently known functions in cancer development is given.

2. Technologies used to study circular RNAs

2.1 Detection of circular RNAs

The widespread abundance of circRNAs remained undiscovered until the advancement of next-generation sequencing (NGS) technology. Due to its unique circular structure, lack of 5'-3' polarity, and lack of a poly-A tail, it's a challenge to accurately detect and quantify circRNAs. At present, the commonly used circRNA detection methods include, Northern blotting, reverse transcription-quantitative PCR (RT-qPCR), RNA fluorescence in situ hybridization (FISH), microarray, and RNA sequencing.

Northern blotting is a method to detect RNA transcripts after size separation by electrophoresis and transfer to a membrane. The circRNA of interest can be detected by hybridization with probes complementary to their nucleotide sequence. The probes used for circRNA detection span the back-splice junction (BSJ) regions or target a sequence being present both in the circular and linear transcripts in combination with a RNase R digestion to degrade the linear transcripts. Northern blotting has been the gold standard for detecting RNA transcripts, but its applicability for circRNAs is lower, as it does not allow detection of its size and generation of specific probes can be challenging. In addition, disadvantages such as low sensitivity, requirement of a large amount of RNA, and being time-consuming make the technique often less attractive [7]. Some research groups have improved the sensitivity and reduced the total assay time of northern blots for circRNAs [8].

RT-qPCR is an assay capable of readily quantifying circRNA levels and is more sensitive than hybridization-based approaches [8, 9]. After cDNA synthesis with random primers, PCR amplification of circRNAs can be achieved using divergent primers allowing amplification of the circRNA and not the linear transcript produced from the same locus **(Figure 1)**. RT-qPCR using RNase R treated RNA is commonly used to further support the circular shape of the circRNA of interest. Sequencing of the PCR products will allow validation of the correctly amplified region. This technique is highly sensitive, specific, fast, and relatively straightforward. However, the potential rolling circle amplification of circRNA molecules during reverse transcription (RT) may lead to an overestimation of its abundance compared to linear RNAs [10]. Droplet digital RT-PCR (ddRT-PCR) is an alternative very sensitive technology to absolutely quantify circRNAs that eliminates the effect of rolling RT products due to the principles of sample partitioning and Poisson statistics. Compared to RT-qPCR, ddRT-PCR has higher sensitivity and accuracy and great potential to be implemented in clinical practice [11].

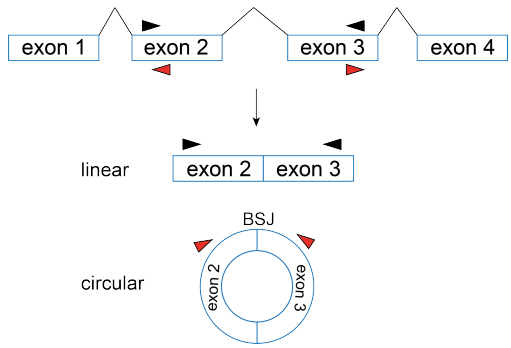

Figure 1. Differences in primer design between linear and circular RNAs. Convergent primers indicated by black arrows are designed to amplify the linear RNAs formed by alternative splicing. Divergent primers indicated by red arrows designed to span the back-splice junction can specifically amplify the circRNAs, but not the counterpart linear RNAs.

RNA-FISH is an effective method for nonquantitative detection of circRNAs in specific cell types or even to define its subcellular localization [8, 12]. Short fluorescently labeled DNA probes specific for the BSJ are hybridized to its target circRNAs on tissue sections or cytospots. After hybridization, fluorescent signals are amplified and visualized with a fluorescence microscope. The need for specific probes targeting the BSJ sequence restricts the probe design and can make the design of effective circRNA probes challenging.

Microarray is an effective high-throughput tool that can be used to comprehensively detect a wide range of circRNAs in multiple samples [13]. A RNase R pretreatment step to remove linear RNA and thereby enrich for circRNAs is required. CircRNAs are reverse-transcribed into fluorescently labeled complementary DNA (cDNA) using random primers. The labeled cDNA is hybridized onto a microarray containing BSJ specific probes. Based on the sequence and GC-content it might be challenging to design effective probes for all circRNAs. In addition, pre-knowledge of the circRNA sequence is required, so this technique does not allow detection of novel circRNAs. Moreover, background hybridization limits the accuracy of expression measurements, particularly for circRNA transcripts present at low abundance. The rolling circle amplification in the reverse transcription step may also lead to overestimation of the abundance of circRNAs. Nevertheless, microarray can be a useful tool for circRNA profiling with a high detection efficiency [13].

RNA sequencing is another high-throughput method, widely used to detect known and novel circRNAs. To increase the efficiency of the RNA-seq for circRNA detection, total RNA is subjected to a ribosomal RNA depletion step, which can be followed by a RNase R treatment to reduce linear RNAs. After sequencing, all reads are aligned to a

reference genome, and non-mapped reads are collected and mapped again to identify split reads indicative of a BSJ event of a circRNA [14]. Several algorithms have been developed for circRNA prediction, such as circRNA_finder, find_circ, CIRI, CIRCexplorer, and others [15]. To reduce false-positive rate for the identification of circRNAs, often several algorithms are combined in combination with a minimal BSJ read count criterium. The false-positive rate for the identification of circRNAs is much lower when RNase R treated samples are used instead of non-RNase R treated RNA samples [15]. An advantage of performing RNA-seq on ribosomal RNA depleted total RNA without RNase treatment, is that in addition to circRNAs, sequencing data will also include expression levels of linear transcripts. This allows guilt-by-association studies and comparison of the levels of circRNAs relative to the levels of the linear RNA of the same transcript. Novel and differentially expressed circRNA candidates can be individually validated and confirmed in subsequent experiments. As indicated, circRNA identification by RNA-seq relies on detection of short reads spanning BSJ that are unique to the circRNA. This means that the internal composition and sequence of circRNAs can at best be in part inferred based on the BSJ sequence. Recently, different long-read nanopore sequencing approaches have been described to analyze the complete structure of specific circRNAs or even circRNAs on a genome wide scale [16].

2.2 Gain- and loss-of-function of circular RNAs

Gain-of-function studies can be performed using overexpression vectors that allow transcription and circRNA formation through back splicing. CircRNA loop formation mostly depends on the complementary pairing of repeats in introns [17] or the interaction with specific RNA-binding proteins [18]. Therefore, canonical splice sites and inverted complementary sequences flanking the circularizing exons, such as Alu elements [17], have been used to design circular RNA overexpression vectors. As some RNA-binding proteins, i.e. QKI [18], NF90/NF110 [19], and HNRNPL [20], can promote circRNA formation, inclusion of specific RBP binding domains in the flanking introns can improve efficiency of RNA circularization. For example, the pcDNA3.1(+) ZKSCAN1 vector contains minimal upstream and downstream intron short-repeat-sequences that were originally shown to support the production of circZKSCAN1 [17]. A revised circRNA overexpression vector called circR (pcDNA3.1 His C vector backbone of 6.3 kb) contains ~800 bp upstream intron of MLLT3, and on the other site of circRNA insert the reverse complementary sequence of this ~800 bp region. The circRNA insert consists of the circRNA sequence together with about 200 bp endogenous flanking genomic sequence [21]. Circularization of the Drosophila laccase2 gene is regulated by both intronic repeats and trans-acting splicing factors hnRNPs and SR proteins. The plasmids

including laccase2 intronic repeats facilitate efficient circular RNA production in mammalian cells [22]. These vectors have been instrumental in studying the function of circRNAs.

Loss-of-function studies are based on specific targeting of the circRNAs without affecting the linear transcripts of the same gene. This can be achieved by RNA interference (RNAi)-based methods including small interfering RNAs (siRNAs) and short hairpin RNAs (shRNAs) designed to specifically target the BSJ sites. While siRNA-based knockdown is especially suitable for short-term studies, shRNA vectors allow for stable transfection of cells resulting in an effective knockdown over a longer period [23]. Many tools can be used for circRNA siRNA/shRNA design, such as CircInteractome (https://circinteractome.nia.nih.gov/) [24]. Recently, a new circRNA knockdown system, RfxCas13d–BSJ-gRNA, based on the principle that type VI CRISPR effectors can be directed to cleave single-stranded (ss) RNA targets. By designing guides that target the BSJ, this will allow effective knockdown of circRNAs without disturbing linear counterparts [25]. In addition to targeting the circRNA at the transcript level, CRISPR/Cas proteins have also been used to knock out circRNAs at the DNA level. This can be achieved by CRISPR-Cas9 mediated genome editing aiming to remove the entire circle-forming exon, which may affect linear transcripts transcription [26] or to disrupt intronic complementary sequences (ICS) flanking the circRNA that drive circRNA formation [27]. Combining base editors with distinct CRISPR/Cas proteins can be applied to change a few bases at splice sites that are predominantly used for back splicing. This approach results in knockout of the circRNA at the genomic level without affecting their cognate linear RNAs [28].

2.3 Identification of molecules binding to circular RNAs

RNA pulldown experiments are one of the main approaches to identify interaction partners of circRNAs. The main steps of a pulldown experiment are the design of a biotinylated antisense oligonucleotide (ASO) targeting the BSJ region, incubation with cell lysates, and pulldown of the ASO-circRNA-interaction partner complexes. Interaction partners can be DNA, RNA molecules, or proteins. DNA interacting partners can be identified by pulldown combined with chromatin immunoprecipitation (ChIP) assay. After cross-linking, circRNAs are reverse transcribed in situ using circRNA-specific antisense primers and biotin-dCTP. After sonication, biotin-cDNA chromatin complexes are pulled down with streptavidin beads. The circRNA-interacting target DNAs are eluted and subjected to DNA sequencing [29]. RNA molecules can be analyzed after pulldown by RT-qPCR in case the binding partners are known or (small) RNA sequencing in case the RNA binding partners are unknown. In case of circRNA-

miRNA interactions, these interactions could be further confirmed using AGO2 immunoprecipitation (AG02-IP) [30]. Protein binding partners can be detected after pulldown by western blotting or mass spectrometry. Results of these experiments can be validated by applying RNA immunoprecipitation (RIP) using an antibody directed against the protein that was pulled down. After separation and purification, presence of the circRNA can be verified by RT-qPCR [12]. ASOs targeting circRNA BSJ may partially bind to linear RNA transcripts nonspecifically, so appropriate negative and positive controls are necessary [31]. For low abundant circRNAs, overexpressing of the circRNA of interest may be required prior to the RNA pulldown procedure [12].

3. B-cell lymphoma

Lymphomas comprise a group of malignancies that are subdivided based on their cellular origin in B- T- or NK-cell lymphomas [32]. Almost 95% of all diagnosed lymphomas are of B-cell origin [33]. B cells mature in the bone marrow and express B-cell receptors (BCRs) on their cell membrane. Naive B cells migrate from the circulation to lymphoid organs once they encounter a cognate antigen and form germinal centers (GCs). The germinal center (GC) represents the site of sustained B-cell proliferation and differentiation. Most B-cell lymphomas originate from GC B cells [34, 35]. These B-cell lymphomas are subdivided into Hodgkin lymphoma (HL) and non-Hodgkin lymphoma (NHL). NHL includes multiple distinct subtypes, with characteristic clinical behavior, morphology, and genetic profiles. Approximately 85-90% of NHL cases are derived from mature B-cell lymphomas, including diffuse large B-cell lymphoma (DLBCL), mantle cell lymphoma (MCL), Burkitt lymphoma (BL), marginal zone lymphoma (MZL), and follicular lymphoma (FL). Among these, DLBCL is the most common subtype with 25-35% of all NHL cases in developed countries [36].

A characteristic feature of HL is the minority of tumor cells in an environment containing a variety of immune cells. HL is subclassified into classical HL (cHL) and nodular lymphocyte-predominant HL (NLPHL) based on immunophenotype and morphology. CHL accounts for over 90% of HL and commonly presents in adolescents and young adults, whereas NLPHL (about 5%) is rare, occurring at a median age of 37 years, and is more common in men [37-39]. The malignant cells in cHL are known as Hodgkin and Reed-Sternberg cells (HRS cells), which have lost their B-cell phenotype but regularly express CD15 and CD30. About 25-40% of cHL cases are positive for Epstein-Barr virus (EBV). The pathogenesis of cHL has been linked to constitutive activation of the nuclear factor κB (NF-κB) and JAK/STAT signaling pathways, caused by genetic aberrations and in case of EBV+ cHL by expression of EBV proteins LMP1 and 2. The malignant cells in NLPHL are called lymphocyte-predominant (LP) cells, these cells

originate from GC B-cells and express both CD45 and CD20. In 40% of NLP cases, SOCS1 is inactivated by somatic mutations in LP cells [36-38].

DLBCL comprises a heterogeneous group of lymphomas characterized by large neoplastic B cells. DLBCL is common in older individuals, with a median age at diagnostic in the mid-60s [40]. DLBCL can be divided into distinct subtypes based on their site of presentation. However, the majority of the DLBCL cases cannot be subdivided and are referred to as DLBCL not otherwise specified (DLBCL-NOS). Gene expression studies have indicated the presence of three distinct DLBCL-NOS subgroups based on their expression profile resembling GC B cells (GCB-DLBCL), activated B cells (ABC-DLBCL), and unclassifiable cases [40]. GCB-DLBCL has a better prognosis (5-year PFS rates 70-80%) than ABC-DLBCL (5-year PFS rates 40-50%) [36]. GCB-DLBCL is associated with PI3K signaling, whereas ABC-DLBCL is characterized by chronic BCR signaling and activation of NF-κB [40]. MYC rearrangements occur in approximately 10% to 15% of DLBCL and are more common in GCB-DLBCL. Cases with MYC and BCL2 and/or BCL6 rearrangements referred to as double- or triple-hit (DH/TH) lymphoma, have a poor outcome. Patients with MYC-DH/TH that involve BCL2 are almost exclusively found in GCB-DLBCL subgroup. The adverse prognostic impact of MYC rearrangements is evident only in patients with a concurrent rearrangement of BCL2 and/or BCL6 and an immunoglobulin (IG) partner [41].

BL is a highly aggressive and fast-growing lymphoma subtype. Three main epidemiological variants of BL are recognized, i.e. endemic, sporadic, and immunodeficiency-associated variants [36]. Endemic BL is the most common childhood malignancy in malaria-endemic areas. The peak incidence of endemic BL in children is from 4 to 7 years old, and boys accounted for more than 60% of all cases. 95% of endemic BL cases are EBV-positive. Sporadic BL is more common in western Europe and the USA and approximately 15-30% of the sporadic cases are EBV-positive in the United States [42]. Immunodeficiency-associated BL is usually associated with HIV infection [36]. BL is characterized by a chromosomal translocation affecting the MYC gene locus at 8q24 and one of the Ig loci. The most commonly observed translocation is t(8;14)(q24;q32), involving MYC and the immunoglobulin heavy (IGH) chain locus, which accounts for 70–80% of cases [43]. Translocation affecting the immunoglobulin light genes, i.e. t(2;8)(p12;q24), and t(8;22)(q24;q11) are less common [36].

4. Circular RNAs in B-cell lymphoma

Although some studies report on the expression and function of circRNAs in B-cell lymphomas, current knowledge is limited. For some circRNAs a role has been shown in

B-cell lymphoma proliferation, migration, cell cycle, and apoptosis by acting as miRNA sponges/baits, stabilizing mRNA, or binding to transcriptional regulators. In addition, circRNA has been associated with clinical characteristics and chemotherapy sensitivity. In some studies, their potential role as a biomarker has been investigated (**Figure 1**).

Dahl et al. reported a genome-wide landscape of circRNAs in MCL and multiple myeloma (MM) cell lines using RNA-seq. Based on the results, 52 unique circRNAs were quantified in patient samples and cell lines from B-cell malignancies using NanoString technology [44]. In addition, they also provide a genome-wide profiling of circRNA expression patterns in an exploratory cohort consisting of 14 diagnostic MCL samples and six healthy controls. Based on circRNA profiles and evaluation of individual candidates as potential biomarkers, nine circRNAs were selected to establish a circRNA-based risk signature (circSCORE). This signature was significantly associated with outcome in two MCL cohorts (n=74 and n=89) [45]. Low circ_RAB11FIP1 was significantly associated with a shorter median time to progression in cases with TP53 mutations, indicating a possible functional link between circ_RAB11FIP1 and TP53 in the pathogenesis of MCL [46]. MCL cases with a high expression of circ_cgga162 had a significantly shorter progression-free and overall survival. Multivariate Cox regression analysis showed that circ_cgga162 expression was an independent prognostic factor for survival [47]. Circ_CDYL was detected at high levels in plasma of MCL patients and knockdown circ_CDYL inhibited MCL cell proliferation [48].

In BL, circ_NSUN2 was shown to be highly expressed and it promoted proliferation, migration, and invasion by binding to the 3'UTR of HMGA1 and thereby stabilizing HMGA1 transcripts. High expression of circ_NSUN2 was achieved by binding of the transcription factor NRF1 to the circ_NSUN2 promoter [49]. Expression of circ_ZDHHC11 was increased in BL cells as compared to normal GC B cells. Circ_ZDHHC11 contains a region with 18 tandemly repeated miR-150 binding sites. This, together with the observation that circ_ZDHHC11 is strongly enriched in the AGO2-IP fraction of BL cells that overexpress miR-150, suggested that circ_ZDHHC11 could function as an endogenous sponge for miR-150 [50]. Simultaneous knockdown of circ_ZDHHC11 and the two linear ZDHHC11 transcripts strongly reduced growth of BL cells, although the underlying mechanism remains unknown.

In DLBCL, circ_0003645 expression was significantly upregulated and it facilitated viability, cell cycle, and glycolysis of DLBCL cells by binding miR-335-5p and thereby releasing miR-335-5p target gene NFIB from miRNA regulation [51]. Circ_OTUD7A was highly expressed in DLBCL and promoted cell proliferation and metastasis, and inhibited cell cycle arrest and apoptosis by binding miR-431-5p and thereby preventing

regulation of FOXP1 by miR-431-5p [52]. Overexpression of circ_CFL1 enhanced proliferation and migration of DLBCL cells by binding miR-107 and thereby inducing expression of its target gene HMGB1 [53]. Circ_APC was significantly downregulated in DLBCL tissues and cell lines, and circ_APC plasma levels were reduced in patients. Circ_APC transcripts were evenly distributed over cytoplasm and nucleus, suggesting a dual functionality. By sponging miR-888, cytoplasmic circ_APC increased APC protein levels, while nuclear circ_APC increased APC expression by binding to the APC promoter and inducing recruitment of the DNA demethylase TET1. Thus, low circ_APC levels resulted in low APC levels and this facilitated growth of DLBCL cells. Analysis of 80 DLBCL cases showed that circ_APC potential is an effective diagnostic and prognostic biomarker for DLBCL [54]. In addition to being regulated by circ_APC, APC expression can also be controlled by circ_EAF2 by binding of circ_EAF2 to the EBV-encoded miR-BART19-3p. This prevents binding of miR-BART19-3p to the 3'UTR of APC, leading to upregulation of APC in EBV+ DLBCL. Circ_EAF2 was downregulated in EBV+ DLBCL tissue and serum samples and its levels were associated with prognosis. Ectopic expression of circ_EAF2 in DLBCL enhanced apoptosis and attenuated proliferation [55].

Long non-coding RNA PVT1 (lncPVT1), also known as Plasmacytoma Variant Translocation 1, is located at 8q24 53 kb downstream of MYC. LncPVT1 often co-amplifies with MYC and is a target of chromosomal breaks. LncPVT1 and one of its circRNA transcripts circ_PVT1 (hsa_circ_00001821) are highly expressed in many different cancers and are involved in the pathogenesis of multiple malignancies through miRNA binding or stabilization of c-Myc protein [56-59]. In B-cell lymphoma, lncPVT1 predicted poor prognosis in patients with DLBCL and facilitated DLBCL development via binding miR-34b-5p [60, 61]. Knockdown of lncPVT1 inhibited proliferation of BL by blocking cell cycle progression [62]. However, nothing is known about the role of circ_PVT1 in B-cell lymphoma.

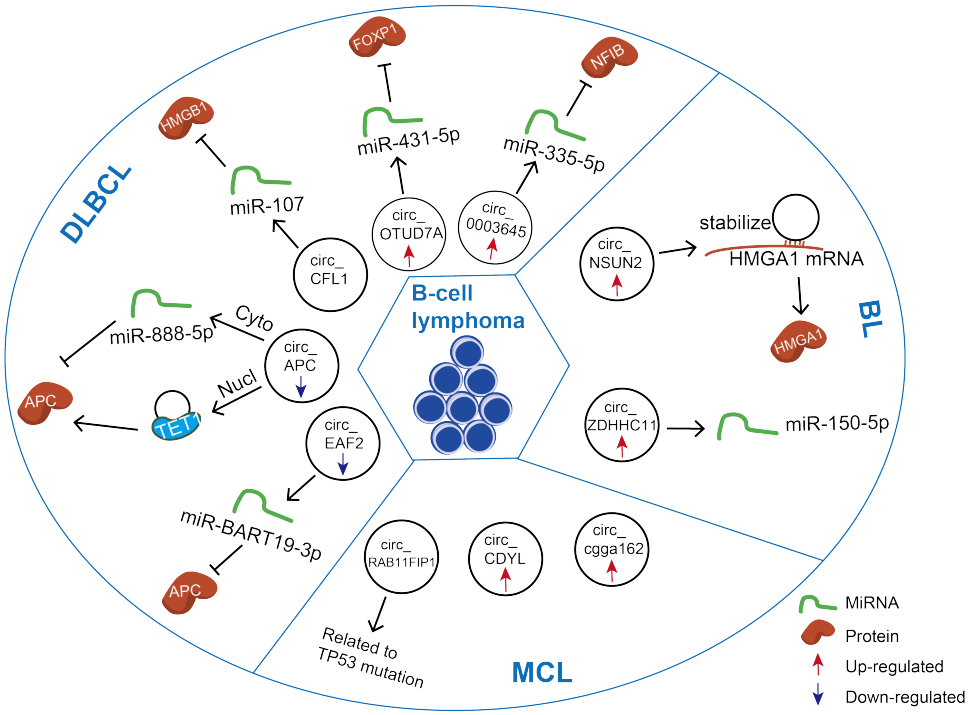

Figure 2. Summary of current knowledge on circRNAs in B-cell lymphoma. A limited set of circRNAs including circ_RAB11FIP1, circ_CDYL, and circ_cgga162 are reported to be diagnostic and prognostic biomarkers in MCL. Functional and mechanistic studies have indicated that circRNA can bind and stabilize mRNA by binding mRNA 3'UTR or serve as a miRNA sponge in BL. In DLBCL, circRNA can bind to miRNAs in the cytoplasm or promotors and recruit transcription regulation factors in the nucleus. MCL, mantle cell lymphoma; DLBCL, diffuse large B-cell lymphoma; BL, Burkitt lymphoma. Red and blue arrows indicate up- and down-regulation of circRNA in cancer; if there are no arrows, this means the expression in difference in comparison to normal cells is unknown.

5. Breast cancer

Breast cancer accounts for approximately 30% of female cancers worldwide, with a mortality-to-morbidity ratio of 15% [63]. Breast cancer is a clinical and molecularly heterogeneous disease with a lifetime risk for a woman of 1:8-1:10 [64]. Subtypes of breast cancer are defined based on the expression of estrogen receptor (ER), progesterone receptor (PR), Ki-67, and human epidermal growth factor receptor 2 (HER2): luminal A, luminal B, HER2 positive, basal-like, and other special histological types [65].

Luminal A type breast cancer is the most common molecular subtype, accounting for 40-50% of all cases [66]. Luminal A tumors are characterized by expression of ER and

PR, a low expression of Ki-67 (less than 20%), and no HER2 expression [65]. Clinically these tumors have a lower histological grade, lower degree of nuclear pleomorphism, lower mitotic activity, and lower proliferation rate than other types of breast tumors. Luminal A subtype is further characterized by expression of luminal epithelial cytokeratins, other luminal-associated markers including ER1, and genes associated with ER function. Patients with luminal A type breast cancer have the best prognosis with fairly high survival and low recurrence rates and respond well to hormonal therapy (tamoxifen) and chemotherapy [67].

Luminal B tumors accounting for 15% to 20% of all breast cancers have a more aggressive phenotype, higher histological grade, and worse prognosis compared to luminal A tumors. Luminal B subtype is subdivided into two subtypes; luminal B-like HER2 negative (ER positive, HER2 negative and Ki-67 high or PR negative or low) and Luminal B-like HER2 positive (ER positive, HER2 overexpressed or amplified) [65]. Approximately 70% of HER2-positive tumors are assigned to the luminal-B subtype [68]. Compared with luminal-A subtype, patients with luminal B are often diagnosed at a younger age and have a higher recurrence rate and lower survival rates after relapse. Luminal B type breast cancer patients have poor outcomes when treated with hormonotherapy but respond better to neoadjuvant chemotherapy [67].

HER2-positive breast cancer comprises 15-20% of all breast cancers and is defined as HER2 overexpressed or amplified with the absence of ER and PR expression [65, 69]. Except for the high expression of the HER2 gene, expression of other genes located on 17q12 are also increased and together this results in a high proliferative, aggressive phenotype. HER2-positive patients generally have a worse prognosis compared to luminal A and B type breast tumors. Around 30% of HER2-positive breast cancers are ER and PR negative [68]. More than 40% of the HER2-positive patients have TP53 gene mutations [70]. Chemotherapy and HER2-targeted monoclonal antibody trastuzumab significantly improved progression-free survival (PFS) and overall survival (OS) in early and advanced-stage of HER2-positive breast cancer.

Basal-like breast cancer is a particularly aggressive molecular subtype defined by a robust set of basal epithelial markers such as keratin 5 and 17, laminin and fatty acid binding protein 7, and low expression of luminal genes [69]. Most, but not all, basal-like tumors are ER, PR, and HER2 gene negative and known as triple-negative (ductal) breast cancer (TNBC). Approximately 85% of basal-like tumors are TNBC [71]. TNBC represents 15-20% of newly diagnosed breast tumors and mostly occurs in premenopausal young women under the age of 40 [72]. Compared with other subtypes of breast cancer, TNBC is more aggressive and has a higher rate of distant

metastasis and early recurrence. TNBC patients usually relapse within 5 years after surgery, and the mortality rate is 40% within 5 years after diagnosis [73]. The median survival time after metastasis was only 13.3 months [74]. TNBC is insensitive to hormone therapy or targeted therapy due to the absence of ER, PR, and HER2. Therefore, chemotherapy is the main systemic treatment, and the use of neoadjuvant chemotherapy can significantly improve the prognosis of TNBC patients. Germline BRCA1 and BRCA2 gene mutations are identified as high-penetrance breast cancer susceptibility genes for TNBC. The presence of BRCA1/2 germline mutations increases the lifetime risk of breast cancer to about 70%, and approximately 15-20% of TNBC patients are associated with BRCA1/2 germline mutations [75]. Several potentially actionable molecular alterations have been found in TNBC, frequently affecting PI3K/mTOR or RAS/RAF/MEK pathway [76].

6. Circular RNAs in breast cancer

For some circRNAs functional roles have been shown in breast cancer by affecting proliferation, migration, invasion, adhesion, cell cycle, apoptosis, and metastatic potential of breast cancer cells (**Figure 3**).

The nuclear circ_ACTN4 was highly expressed in breast cancer and its expression was associated with advanced clinical stage and poor prognosis. Circ_ACTN4 promoted proliferation, migration, and invasive potential of breast cancer cells. Mechanistically, circ_ACTN4 competitively bound FUBP1 and thereby prevented the interaction between FUBP1 and FIR. This induced MYC transcription and facilitated progression of breast cancer [77]. Circ_SCRIB is formed from an antisense SCRIB transcript and includes sequences that are antisense to parts of exons 8 and 9 and the whole intron 8 of the SCRIB gene. Based on these perfect complementary sequences circ_SCRIB binds to the SCRIB pre-mRNA and blocks splicing of intron 8. This reduced the levels of spliced SCRIB transcripts and SCRIB protein, which caused increased proliferation, migration, invasion, and adhesion [78].

For cytoplasmic circRNAs, the regulatory mechanism related to breast cancer progression involved binding to miRNAs. Elevated circ_BACH2 levels were shown to contribute to cell proliferation, epithelial to mesenchymal transition, and malignant progression of TNBC cells through binding miR-186-5p and miR-548c-3p, thereby releasing CXCR4 from being inhibited by these miRNAs. Furthermore, knockdown of circ_BACH2 repressed formation of lung metastasis *in vivo* [79]. Two circRNAs (circ_0043278 and circ_0006220) derived from the TADA2A gene and circ_NOL10 (circ_0000977) were significantly decreased in TNBC [80]. Downregulation of two

circRNAs from TADA2A was confirmed in breast cancer tissues and cell lines. Decreased expression of circ_0006220 but not circ_0043278 was significantly associated with increased lymphatic metastasis, advanced clinical stage, and poor prognosis in TNBC. Functionally, circ_0006220 inhibited miR-203a-3p activity by binding to miR-203a-3p. Depletion of miR-203a-3p resulted in upregulation of the miR-203a-3p target gene SOCS3, and this suppressed cellular proliferation, clonogenicity, migration, and invasion [80]. A further analysis of the role of circ_NOL10 in TNBC has been described in **Chapter 6**.

The cytoplasmic circ_0068631 was upregulated in breast cancer and was shown to stabilize c-Myc transcripts by recruiting EIF4A3 which led to enhanced c-Myc protein levels. In this way, circ_0068631 promoted proliferation and migration of breast cancer cells [81]. Circ_CREIT was significantly downregulated in chemo-resistant TNBC and its levels were associated with poor prognosis. Ectopic overexpression of circ_CREIT enhanced chemosensitivity and chemotherapy-induced apoptosis. Mechanistically, circ_CREIT acted as a scaffold to facilitate the interaction between PKR and the E3 ligase HACE1, and this promoted proteasomal degradation of PKR protein via K48-linked polyubiquitylation. This resulted in suppression of eIF2α phosphorylation and assembly of stress granules [82].

Circ_EIF6 encoding for a 224-amino acid peptide named EIF6-224aa promoted proliferation and metastasis of TNBC cells *in vitro* and *in vivo*. This was achieved by directly interacting with the MYH9 oncogene and reducing its degradation by inhibiting the ubiquitin-proteasome pathway. In addition, high expression of circ_EIF6 was associated with poor overall survival in TNBC patients, and circ_EIF6 expression was associated with histological grade and distant metastasis [83]. Circ_HER2, formed upon back splicing of exons 3-7 of the HER2 gene, levels were enriched in part of TNBCs, and the expression of circ_HER2 was negatively correlated with the overall survival of TNBC patients. Circ_HER2 encoded a novel protein called HER2-103 that enhanced proliferation and invasion of TNBC cells by promoting EGFR/HER3 interaction and activation [84].

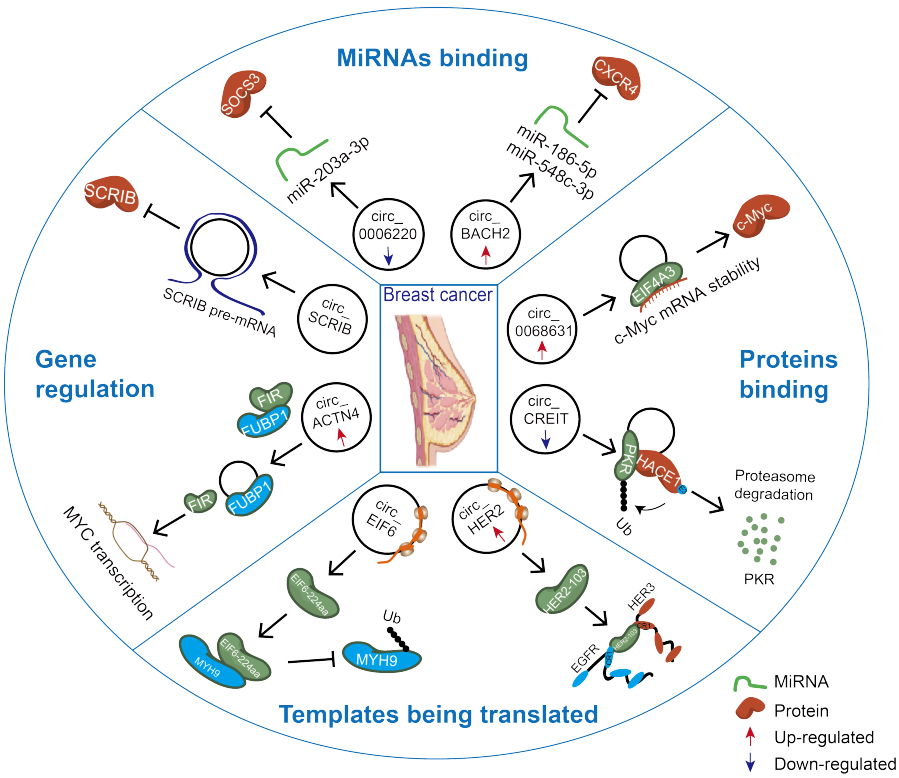

Figure 3. Overview of circRNA action modes involved in breast cancer pathogenesis. CircRNAs in the nucleus can participate in gene regulation by regulating gene transcription or pre-mRNA splicing. CircRNAs in the cytoplasm regulate breast cancer pathogenesis mainly by binding to miRNAs, proteins, or being translated to peptides or proteins. Red and blue arrows indicate up- and downregulation of circRNA in breast cancer.

7. Scope of thesis

As a relatively new field of research, current knowledge on the expression and functions of circRNAs in normal cells and disease conditions is limited. Functional studies focus amongst others on the elucidation of the pivotal roles of circRNAs in general and elucidating their working mechanisms. In addition, expression patterns and functional involvement of circRNAs have been studied in relation to cancer development and progression. Expression profiling studies focus on their potential roles as diagnostic, prognostic, and predictive biomarkers, while functional studies focus on their mode of action and explore their potential as therapeutic targets. At this moment, it is essential to gain a deeper understanding of the physiological and pathophysiological roles of circRNA in cancer. This study aimed to explore the role of

circRNAs including the expression pattern, functional mechanisms, and regulatory roles in B-cell lymphoma and breast cancer.

In **Chapter 2**, we provide an introduction to circRNAs including biogenesis, the role of RNA-binding proteins in this process, different modes of action, and current knowledge on the molecular regulatory mechanisms of circRNAs in human cancers. In **Chapter 3**, we characterized the circRNA landscape of three main B-cell lymphoma subtypes and normal B-cells using RNA sequencing. Circular shape of the identified circRNAs was validated by RNase R treatment and differential expression patterns were confirmed in an extended B-cell lymphoma cell line panel. In **Chapter 4**, we studied the role of circ_PVT1, a circular RNA derived from the PVT1 locus. PVT1 is located 53 kb downstream of MYC and a possible target of amplification and chromosomal breaks in part of lymphoma cases with chromosome 8 aberrations. We focused on the expression patterns of both linear and circular PVT1, effects of circ_PVT1 inhibition and overexpression on proliferation and the possible mode of action in B-cell lymphoma cell lines. As a control, we also analyzed the function of linear PVT1 transcripts. In **Chapter 5**, we focused on the relevance of circ_ZDHHC11 and its miR-150 binding site region for the growth of BL cells. We previously showed that circ_ZDHHC11 has an 18 miR-150 binding site region, the same as other ZDHHC11 transcripts and is enriched in AGO2-IP fraction upon miR-150 overexpression [50]. In this study, we investigated the effects of knockdown and overexpression of circ_ZDHHC11 in unmodified BL cells and in miR-150 binding site knock-out BL cells. In **Chapter 6**, we explored the expression pattern, diagnostic and clinical value of circ_NOL10 in breast cancer progression and metastasis. This candidate was selected based on previous circRNA microarray analysis data, showing that circ_NOL10 was downregulated in breast cancer [80]. In addition, we studied the role and regulatory pathway of circ_NOL10 in breast cancer, and the regulatory mechanism of circ_NOL10 formation. In **Chapter 7**, we summarized and discussed our results and offered perspectives for future research.

Reference

1. Sanger, H.L., et al., *Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures.* Proc Natl Acad Sci U S A, 1976. **73**(11): p. 3852-6.
2. Nigro, J.M., et al., *Scrambled exons.* Cell, 1991. **64**(3): p. 607-13.
3. Chen, I., C.Y. Chen, and T.J. Chuang, *Biogenesis, identification, and function of exonic circular RNAs.* Wiley Interdiscip Rev RNA, 2015. **6**(5): p. 563-79.
4. Rybak-Wolf, A., et al., *Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed.* Mol Cell, 2015. **58**(5): p. 870-85.
5. Salzman, J., *Circular RNA Expression: Its Potential Regulation and Function.* Trends Genet, 2016. **32**(5): p. 309-316.
6. Patop, I.L., S. Wüst, and S. Kadener, *Past, present, and future of circRNAs.* Embo j, 2019. **38**(16): p. e100836.
7. D'Ambra, E. and M. Morlando, *Study of Circular RNA Expression by Nonradioactive Northern Blot Procedure.* Methods Mol Biol, 2021. **2348**: p. 371-383.
8. Mi, Z., et al., *Circular RNA detection methods: A minireview.* Talanta, 2022. **238**(Pt 2): p. 123066.
9. Panda, A.C. and M. Gorospe, *Detection and Analysis of Circular RNAs by RT-PCR.* Bio Protoc, 2018. **8**(6).
10. Cocquet, J., et al., *Reverse transcriptase template switching and false alternative transcripts.* Genomics, 2006. **88**(1): p. 127-31.
11. Sedlak, R.H., J. Kuypers, and K.R. Jerome, *A multiplexed droplet digital PCR assay performs better than qPCR on inhibition prone samples.* Diagn Microbiol Infect Dis, 2014. **80**(4): p. 285-6.
12. Pandey, P.R., et al., *Methods for analysis of circular RNAs.* Wiley Interdiscip Rev RNA, 2020. **11**(1): p. e1566.
13. Li, S., et al., *Microarray is an efficient tool for circRNA profiling.* Brief Bioinform, 2019. **20**(4): p. 1420-1433.
14. Jakobi, T. and C. Dieterich, *Computational approaches for circular RNA analysis.* Wiley Interdiscip Rev RNA, 2019. **10**(3): p. e1528.
15. Hansen, T.B., et al., *Comparison of circular RNA prediction tools.* Nucleic Acids Res, 2016. **44**(6): p. e58.
16. Rahimi, K., et al., *Nanopore long-read sequencing of circRNAs.* Methods, 2021. **196**: p. 23-29.
17. Liang, D. and J.E. Wilusz, *Short intronic repeat sequences facilitate circular RNA production.* Genes Dev, 2014. **28**(20): p. 2233-47.
18. Conn, S.J., et al., *The RNA binding protein quaking regulates formation of circRNAs.* Cell, 2015. **160**(6): p. 1125-34.
19. Li, X., et al., *Coordinated circRNA Biogenesis and Function with NF90/NF110 in Viral Infection.* Mol Cell, 2017. **67**(2): p. 214-227.e7.
20. Fei, T., et al., *Genome-wide CRISPR screen identifies HNRNPL as a prostate cancer dependency regulating RNA splicing.* Proc Natl Acad Sci U S A, 2017. **114**(26): p. E5207-e5215.
21. Liu, D., et al., *A Highly Efficient Strategy for Overexpressing circRNAs.* Methods Mol Biol, 2018. **1724**: p. 97-105.
22. Kramer, M.C., et al., *Combinatorial control of Drosophila circular RNA expression by intronic repeats, hnRNPs, and SR proteins.* Genes Dev, 2015. **29**(20): p. 2168-82.
23. Liu, Y.P., N.C. Schopman, and B. Berkhout, *Dicer-independent processing of short hairpin RNAs.* Nucleic Acids Res, 2013. **41**(6): p. 3723-33.
24. Dudekula, D.B., et al., *CircInteractome: A web tool for exploring circular RNAs and their interacting proteins and microRNAs.* RNA Biol, 2016. **13**(1): p. 34-42.
25. Li, S., et al., *Screening for functional circular RNAs using the CRISPR-Cas13 system.* Nat Methods, 2021. **18**(1): p. 51-59.
26. Piwecka, M., et al., *Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function.* Science, 2017. **357**(6357).
27. Zhang, Y., et al., *The Biogenesis of Nascent Circular RNAs.* Cell Rep, 2016. **15**(3): p. 611-624.
28. Gao, X., et al., *Knockout of circRNAs by base editing back-splice sites of circularized exons.* Genome Biol, 2022. **23**(1): p. 16.
29. Chen, N., et al., *A novel FLI1 exonic circular RNA promotes metastasis in breast cancer by coordinately regulating TET1 and DNMT1.* Genome Biol, 2018. **19**(1): p. 218.
30. Yang, J.H., P.R. Pandey, and M. Gorospe, *Identification of circRNA-Interacting Proteins by Affinity Pulldown.* Methods Mol Biol, 2021. **2372**: p. 193-202.
31. Du, W.W., et al., *Identifying and Characterizing circRNA-Protein Interaction.* Theranostics, 2017. **7**(17): p. 4183-4191.
32. Armitage, J.O., et al., *Non-Hodgkin lymphoma.* Lancet, 2017. **390**(10091): p. 298-310.
33. Küppers, R., *Mechanisms of B-cell lymphoma pathogenesis.* Nat Rev Cancer, 2005. **5**(4): p. 251-62.
34. Basso, K. and R. Dalla-Favera, *Germinal centres and B cell lymphomagenesis.* Nat Rev Immunol, 2015. **15**(3): p. 172-84.

35. Koues, O.I., E.M. Oltz, and J.E. Payton, *Short-Circuiting Gene Regulatory Networks: Origins of B Cell Lymphoma.* Trends Genet, 2015. **31**(12): p. 720-731.
36. Swerdlow SH, C.E., Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues.* 2017. **2**.
37. Küppers, R., A. Engert, and M.L. Hansmann, *Hodgkin lymphoma.* J Clin Invest, 2012. **122**(10): p. 3439-47.
38. Shanbhag, S. and R.F. Ambinder, *Hodgkin lymphoma: A review and update on recent progress.* CA Cancer J Clin, 2018. **68**(2): p. 116-132.
39. Eichenauer, D.A. and A. Engert, *Nodular lymphocyte-predominant Hodgkin lymphoma: a unique disease deserving unique management.* Hematology Am Soc Hematol Educ Program, 2017. **2017**(1): p. 324-328.
40. Sehn, L.H. and G. Salles, *Diffuse Large B-Cell Lymphoma.* N Engl J Med, 2021. **384**(9): p. 842-858.
41. Rosenwald, A., et al., *Prognostic Significance of MYC Rearrangement and Translocation Partner in Diffuse Large B-Cell Lymphoma: A Study by the Lunenburg Lymphoma Biomarker Consortium.* J Clin Oncol, 2019. **37**(35): p. 3359-3368.
42. Klumb, C.E., et al., *Geographic variation in Epstein-Barr virus-associated Burkitt's lymphoma in children from Brazil.* Int J Cancer, 2004. **108**(1): p. 66-70.
43. Molyneux, E.M., et al., *Burkitt's lymphoma.* Lancet, 2012. **379**(9822): p. 1234-44.
44. Dahl, M., et al., *Enzyme-free digital counting of endogenous circular RNA molecules in B-cell malignancies.* Lab Invest, 2018. **98**(12): p. 1657-1669.
45. Dahl, M., et al., *Expression patterns and prognostic potential of circular RNAs in mantle cell lymphoma: a study of younger patients from the MCL2 and MCL3 clinical trials.* Leukemia, 2022. **36**(1): p. 177-188.
46. Dahl, M., et al., *A Circular RNA Molecule, circRAB11FIP1, Is Associated with TP53 Mutations and Is of Potential Prognostic and Functional Significance in Mantle Cell Lymphoma: Data from the Nordic MCL2 and MCL3 Studies.* Blood, 2019. **134**(Supplement_1): p. 1495-1495.
47. Li, Y., et al., *[Relationship of Expression of Circ_cgga162 with the Prognosis of Patients with Mantle Cell Lymphoma].* Zhongguo Shi Yan Xue Ye Xue Za Zhi, 2020. **28**(3): p. 876-880.
48. Mei, M., et al., *CircCDYL Serves as a New Biomarker in Mantle Cell Lymphoma and Promotes Cell Proliferation.* Cancer Manag Res, 2019. **11**: p. 10215-10221.
49. Wang, L., et al., *NRF1-regulated CircNSUN2 promotes lymphoma progression through activating Wnt signaling pathway via stabilizing HMGA1.* Cell Cycle, 2021. **20**(9): p. 819-828.
50. Dzikiewicz-Krawczyk, A., et al., *ZDHHC11 and ZDHHC11B are critical novel components of the oncogenic MYC-miR-150-MYB network in Burkitt lymphoma.* Leukemia, 2017. **31**(6): p. 1470-1473.
51. Lin, D., et al., *Circ_0003645 serves as miR-335-5p sponge to promote the biological process of diffuse large B-cell lymphoma by upregulating NFIB.* Autoimmunity, 2022. **55**(2): p. 127-135.
52. Liu, W., et al., *CircRNA_OTUD7A upregulates FOXP1 expression to facilitate the progression of diffuse large B-cell lymphoma via acting as a sponge of miR-431-5p.* Genes Genomics, 2021. **43**(6): p. 653-667.
53. Chen, X., X. Xie, and W. Zhou, *CircCFL1/MiR-107 Axis Targeting HMGB1 Promotes the Malignant Progression of Diffuse Large B-Cell Lymphoma Tumors.* Cancer Manag Res, 2020. **12**: p. 9351-9362.
54. Hu, Y., et al., *A circular RNA from APC inhibits the proliferation of diffuse large B-cell lymphoma by inactivating Wnt/β-catenin signaling via interacting with TET1 and miR-888.* Aging (Albany NY), 2019. **11**(19): p. 8068-8084.
55. Zhao, C.X., et al., *CircEAF2 counteracts Epstein-Barr virus-positive diffuse large B-cell lymphoma progression via miR-BART19-3p/APC/β-catenin axis.* Mol Cancer, 2021. **20**(1): p. 153.
56. Jin, K., et al., *Long non-coding RNA PVT1 interacts with MYC and its downstream molecules to synergistically promote tumorigenesis.* Cell Mol Life Sci, 2019. **76**(21): p. 4275-4289.
57. Li, R., et al., *lncRNA PVT1: a novel oncogene in multiple cancers.* Cell Mol Biol Lett, 2022. **27**(1): p. 84.
58. Mo, Y., et al., *Circular RNA circPVT1 promotes nasopharyngeal carcinoma metastasis via the β-TrCP/c-Myc/SRSF1 positive feedback loop.* Mol Cancer, 2022. **21**(1): p. 192.
59. Palcau, A.C., et al., *CircPVT1: a pivotal circular node intersecting Long Non-Coding-PVT1 and c-MYC oncogenic signals.* Mol Cancer, 2022. **21**(1): p. 33.
60. Yang, R., et al., *Long noncoding RNA PVT1 promotes tumor growth and predicts poor prognosis in patients with diffuse large B-cell lymphoma.* Cancer Commun (Lond), 2020. **40**(10): p. 551-555.
61. Tao, S., et al., *LncRNA PVT1 facilitates DLBCL development via miR-34b-5p/Foxp1 pathway.* Mol Cell Biochem, 2022. **477**(3): p. 951-963.
62. Zheng, C., et al., *Knockdown of long non-coding RNA PVT1 inhibits the proliferation of Raji cells through cell cycle regulation.* Oncol Lett, 2019. **18**(2): p. 1225-1234.
63. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2020.* CA Cancer J Clin, 2020. **70**(1): p. 7-30.
64. Harbeck, N. and M. Gnant, *Breast cancer.* Lancet, 2017. **389**(10074): p. 1134-1150.
65. Goldhirsch, A., et al., *Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013.* Ann Oncol, 2013. **24**(9): p. 2206-23.

66. Tsang, J.Y.S. and G.M. Tse, *Molecular Classification of Breast Cancer.* Adv Anat Pathol, 2020. **27**(1): p. 27-35.
67. Yersal, O. and S. Barutca, *Biological subtypes of breast cancer: Prognostic and therapeutic implications.* World J Clin Oncol, 2014. **5**(3): p. 412-24.
68. Fan, Y., et al., *Clinical features of patients with HER2-positive breast cancer and development of a nomogram for predicting survival.* ESMO Open, 2021. **6**(4): p. 100232.
69. Alluri, P. and L.A. Newman, *Basal-like and triple-negative breast cancers: searching for positives among many negatives.* Surg Oncol Clin N Am, 2014. **23**(3): p. 567-77.
70. Tsutsui, S., et al., *Prognostic significance of the coexpression of p53 protein and c-erbB2 in breast cancer.* Am J Surg, 2003. **185**(2): p. 165-7.
71. Ensenyat-Mendez, M., et al., *Current Triple-Negative Breast Cancer Subtypes: Dissecting the Most Aggressive Form of Breast Cancer.* Front Oncol, 2021. **11**: p. 681476.
72. Morris, G.J., et al., *Differences in breast carcinoma characteristics in newly diagnosed African-American and Caucasian patients: a single-institution compilation compared with the National Cancer Institute's Surveillance, Epidemiology, and End Results database.* Cancer, 2007. **110**(4): p. 876-84.
73. Dent, R., et al., *Triple-negative breast cancer: clinical features and patterns of recurrence.* Clin Cancer Res, 2007. **13**(15 Pt 1): p. 4429-34.
74. Lin, N.U., et al., *Sites of distant recurrence and clinical outcomes in patients with metastatic triple-negative breast cancer: high incidence of central nervous system metastases.* Cancer, 2008. **113**(10): p. 2638-45.
75. Loibl, S., et al., *Breast cancer.* Lancet, 2021. **397**(10286): p. 1750-1769.
76. Bianchini, G., et al., *Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease.* Nat Rev Clin Oncol, 2016. **13**(11): p. 674-690.
77. Wang, X., et al., *The circACTN4 interacts with FUBP1 to promote tumorigenesis and progression of breast cancer by regulating the expression of proto-oncogene MYC.* Mol Cancer, 2021. **20**(1): p. 91.
78. Ma, J., et al., *An antisense circular RNA circSCRIB enhances cancer progression by suppressing parental gene splicing and translation.* Mol Ther, 2021. **29**(9): p. 2754-2768.
79. Wang, X., et al., *Up-regulated circBACH2 contributes to cell proliferation, invasion, and migration of triple-negative breast cancer.* Cell Death Dis, 2021. **12**(5): p. 412.
80. Xu, J.Z., et al., *circTADA2As suppress breast cancer progression and metastasis via targeting miR-203a-3p/SOCS3 axis.* Cell Death Dis, 2019. **10**(3): p. 175.
81. Wang, X., M. Chen, and L. Fang, *hsa_circ_0068631 promotes breast cancer progression through c-Myc by binding to EIF4A3.* Mol Ther Nucleic Acids, 2021. **26**: p. 122-134.
82. Wang, X., et al., *CircRNA-CREIT inhibits stress granule assembly and overcomes doxorubicin resistance in TNBC by destabilizing PKR.* J Hematol Oncol, 2022. **15**(1): p. 122.
83. Li, Y., et al., *circ-EIF6 encodes EIF6-224aa to promote TNBC progression via stabilizing MYH9 and activating the Wnt/beta-catenin pathway.* Mol Ther, 2022. **30**(1): p. 415-430.
84. Li, J., et al., *Circular HER2 RNA positive triple negative breast cancer is sensitive to Pertuzumab.* Mol Cancer, 2020. **19**(1): p. 142.