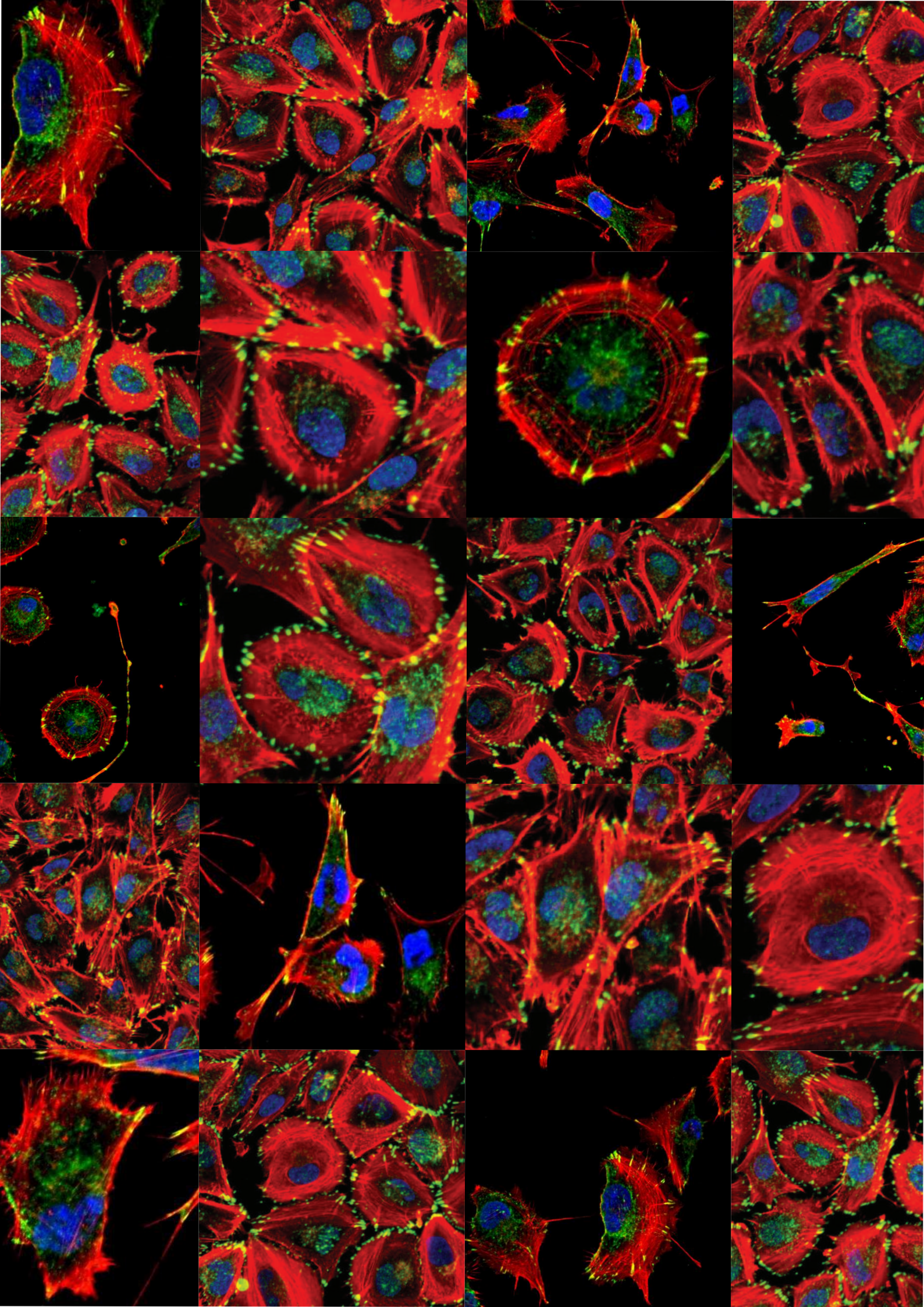
RNA splicing in breast cancer progression

Koedoot, E.



Chapter 1

Introduction, aim and scope of the thesis

Highlights

- Triple-negative breast cancer (TNBC) is a very aggressive breast cancer subtype disproportionally responsible for metastasis formation and related breast cancer deaths
- Tumor cell migration and tumor cell proliferations are essential processes for cancer progression and metastasis formation
- In this thesis, we aimed to better understand the complete signaling landscape driving TNBC cell migration and proliferation and discover new targets to reduce TNBC-related deaths

◄ IN THE PICTURE

Microscopic view of triple-negative breast cancer cells with(out) knockdown of genes involved in cell migration. Cells were stained for the actin cytoskeleton (red), migratory protein p-FAK (green) and nuclei (blue).

◄ IN BEELD

Microscopisch beeld van triple-negatieve borstkankercellen met of zonder uitschakeling van genen die betrokken zijn bij tumorcelmigratie. Cellen werden gekleurd voor het actine cytoskelet (rood), migratie-eiwit p-FAK (groen) en celkernen (blauw).

Chapter 1

Breast cancer is the most prevalent cancer in women with 562,500 and 266,120 estimated new cases Europe[1] and the United States[2] in 2018, respectively. In total 1 out of 8 women will develop breast cancer before the age of 85[3]. Although the 5 year survival rate is 91%[4], breast cancer still has the highest cancer mortality rate in women worldwide[1]. 90% of these deaths are caused by the formation of secondary tumors in distant organs, also called metastases[5]. This thesis focused on unraveling the underlying mechanisms and the discovery of new targets involved in this process of metastasis formation, with the ultimately goal to reduce breast cancer related deaths. In this chapter, the different aspects of breast cancer classification, progression and metastasis formation will be discussed.

Breast cancer classification

Breast cancer is a heterogeneous disease that can arise from various areas in the breast. Based on the cellular origin from which the tumor evolved, breast cancer can be classified as either carcinoma (derived from epithelial cells) or sarcoma (derived from stromal compartments). More than 99% of the breast tumors are carcinomas that can be distributed over three major groups based on pathological characteristics and invasiveness: 1) non-invasive tumors that develop inside the normal ducts, 2) invasive or infiltrating tumors that grow into the surrounding stromal tissue and 3) metastatic tumors that spread to other organs, such as the lungs, brain or liver. Precise distinction between different subtypes is essential for accurate treatment of breast cancer patients. Therefore, next to the histological characterization, breast cancer has also been classified by genetic profiling and hormone receptor (HR) expression levels[6–8]. Based on gene expression profiles, five major molecular subtypes were identified: claudin-low, basal-like, luminal A, luminal B and human epidermal growth factor receptor 2 or HER2-enriched[7,8]. The luminal A and B subtypes are mostly HR positive, depend on HR downstream signaling pathways and can therefore often be treated with anti-hormonal therapy[6]. The HER2-enriched subtype demonstrates overexpression of the HER2 receptor and is treated using anti-HER2 monoclonal antibodies or HER2-tyrosine kinase inhibitors[9–12]. In contrast to the luminal and HER2 subtypes, triple negative breast cancer (TNBC) lacks amplification of HRs such as the estrogen receptor (ER), progesterone receptor (PR) or human epidermal growth factor receptor 2 (HER2)[6]. As a consequence, TNBC does not respond to anti-hormonal therapies and there is not yet a targeted therapy for TNBC available. Therefore, the current treatment strategy for TNBC still consist of conventional surgery and unselective aggressive treatments, such as chemo- and radiotherapy[13]. Although TNBC only accounts for 15-20% of all breast cancer cases[14], it is disproportionally responsible for metastasis formation and related breast cancer deaths compared to the HR positive breast cancer subtypes[15]. Therefore, unraveling the underlying mechanisms involved in TNBC progression is essential to reduce breast cancer induced mortality.

The metastatic cascade

As discussed above, metastases are the underlying cause of death for the majority of breast cancer patients. Metastasis formation is a highly complicated multistep process (Figure 1) that is partly mediated by intrinsic changes in the tumor cells as well as extrinsic effects of the tumor

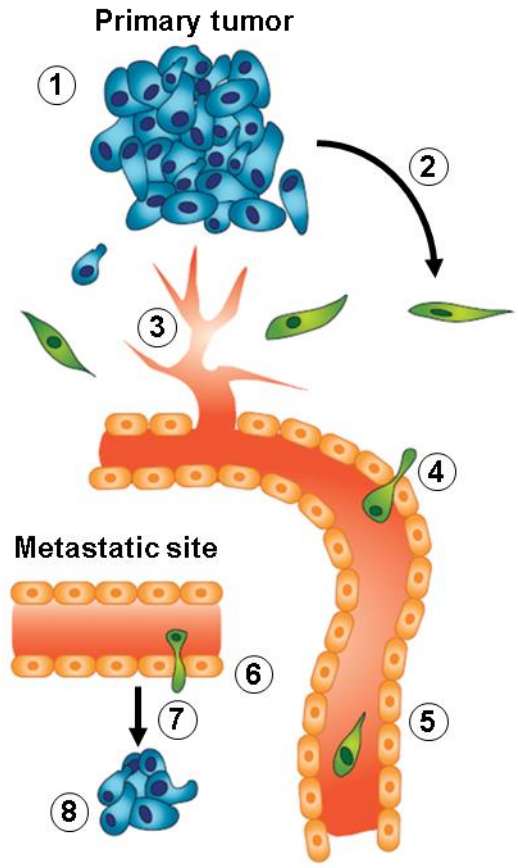


Figure 1. ***Overview of the metastatic cascade.*** After the primary tumor has formed (1), tumor cells can gain invasive properties via epithelial-to-mesenchymal transition (2) and/or inducing angiogenesis (3). Tumor cells invade (4) and travel (5) through the circulatory system, followed by extravasation at distant metastatic sites (6). Next tumor cells need to undergo mesenchymal-to-epithelial transition (7) and start proliferating to finally form macro-metastases (8).

microenvironment[16–20]. After the primary tumor has formed, tumor cells need to gain invasive features to metastasize. These characteristics can be obtained via genomic alterations resulting in elevated levels of oncogenes or downregulation of tumor suppressors[21–23] or stimuli from the tumor microenvironment, such as chemokines and macrophages[24,25]. Altogether, this leads to epithelial-to-mesenchymal transition (EMT), a process in which epithelial cells lose their cell-cell contacts and polarized organization and acquire mesenchymal characteristics resulting in a more migratory and invasive phenotype[26]. At the same time, tumor cells induce the growth of blood and lymphatic vessels into the tumor (also called angiogenesis), which is essential to fuel the cells with nutrients but also provides a way to escape from the primary tumor. Next cancer cells can invade into the blood vessels, followed by survival in the circulatory system and extravasation at distant organs[20]. After extravasation, tumor cells need to undergo mesenchymal-to-epithelial transition (MET), which allows the cells to switch back to the epithelial state. Tumor cells at the metastatic site frequently reside in a dormal state and need interactions with the local microenvironment to start proliferating and form macro-metastases[18].

In the metastatic cascade described above, we can appreciate two processes pivotal for tumors to progress and metastasize: 1) tumor cell migration and 2) tumor cell proliferation. These processes will be discussed in more detail below.

Tumor cell migration

Tumor cell migration is involved in many steps of the metastatic cascade, including local invasion, invading the circulation, extravasation and dissemination to distant tissue[27]. Cell migration starts with the protrusion of the cell membrane driven by tightly regulated actin polymerization at the leading edge of the cell[28,29]. These flat membrane protrusions or lamellipodia generate a force that is necessary to pull the cell forward[30] and regulate the actin organization within the cell[31]. The lamellipodia are stabilized by attachment to the extracellular matrix (ECM) through cell-matrix adhesions[32]. The cell is then pulled forward by actomyosin contraction[30], followed by local disassembly of the cell-matrix adhesions at the rear of the

cell[33,34] thereby completing the cell migration cycle. Next, the cell will start with a new cycle by creating new protrusions at the leading edge of the cell.

The cell-matrix adhesions or focal adhesions are mediated by large dynamic and flexible protein complexes and play a pivotal role during the migration cycle. The loss of focal adhesions at the rear of the cell together with the induction of new adhesions at the leading edge provides physical support while simultaneously allowing the cell to move forward[33,34]. In these complexes, integrins form the critical link between the cell and the ECM. Intracellularly, these integrins are bound to a complex consisting of more than 180 proteins including paxillin, vinculin and focal adhesion kinase, that connect the integrins to the actin cytoskeleton[35,36]. Since these processes are highly dynamic in time and space, tight regulation of the focal adhesion complexes is essential for efficient cell migration[37,38].

Many TNBC cell lines demonstrate increased plasticity, a mesenchymal phenotype and high migratory behavior associated with metastasis formation[24,39–44]. Therefore, uncovering the mechanisms of TNBC cell migration is critical to understand the progression of metastatic disease.

Tumor cell proliferation

Cell proliferation is essential for tumor outgrowth in the primary as well as metastatic tumor and is therefore the most fundamental and most studied cancer cell characteristic. In normal tissue, cell cycle related processes are tightly regulated thereby controlling cell number, tissue architecture, homeostasis and tissue functioning. Cancer cells often hijack these pathways via different ways: 1) self-sufficiency in growth signals, 2) decreased sensitivity to anti-proliferative signaling, 3) limitless replicative potential and 4) resisting apoptosis[21,22]. Often these effects are mediated by increased levels of oncogenes such as HER2, MYC and PI3KCA or loss of function mutations in tumor suppressor genes such as TP53, RB, BRCA1/2 and PTEN[45–47], thereby allowing the cells to keep proliferating. In breast cancer, sustained proliferation is often mediated by overexpression of the HRs, such as the ER, HER2 and/or PR and its corresponding ligands. These ligands can be supplied by the cancer cells itself leading to autocrine proliferative signaling or otherwise, cancer cells can stimulate cells in the tumor microenvironment to produce these growth factors[48,49]. Active HRs then stimulate proliferative pathways such as the mitogen activated protein-kinase (MAPK) and the phosphoinositide 3-kinase (PI3K) pathways[50–52]. However, as explained above, since the TNBC subtype lacks expression of HRs it cannot be treated with HR-specific inhibitors. Therefore, the signaling pathways that control TNBC proliferation have been investigated using genome-wide screens[53,54] and inhibitors for different oncogenic pathways have been developed[55–57]. Unfortunately, often TNBC does not respond to these inhibitors or acquired resistance is developed during treatment[58]. Therefore, there is still a high demand for developing new treatments to efficiently halt TNBC cell proliferation thereby preventing metastasis formation.

Identifying signaling components essential in TNBC migration and proliferation

Although the role of some individual genes in TNBC has previously been described[59–63], the complete signaling pathways driving TNBC cell motility, proliferation and metastasis formation

are less clear. To discover determinants of TNBC metastasis formation, we established a breast cancer cell line panel (46 cell lines) including 18 TNBC cell lines that has been systematically characterized based on morphology, proliferation and migratory behavior (Rogkoti et al, manuscript in preparation). Next to conventional 2D cell cultures, these cell lines can also be used in more advanced 3D gel compositions that better resemble human tissue physiology. Moreover, the introduction of large RNAi libraries allows the identification of potential drug targets at a genome-wide scale, resulting in the discovery of multiple disease signaling components from a single screen. To apply these RNAi libraries in the discovery of TNBC cell motility determinants, we established two automatic imaging-based assays (phagokinetic track (PKT) and random cell migration (RCM) assays) that can be used in screening settings[64,65]. This approach already resulted in the discovery of the splicing factor SRPK1 as a critical determinant in breast cancer metastasis formation[66]. During the splicing reaction, non-coding pre-mRNA regions (or introns) are removed resulting in a transcript that can be translated into a functional protein[67]. Moreover, by including specific exons, splicing can regulate the expression of different isoforms derived from the same gene, providing an extra layer of genetic control[68]. The splicing reaction is catalyzed by a RNA-protein complex (the spliceosome) consisting of 5 small nuclear RNAs (snRNAs) and 244 associated proteins or splicing factors, amongst which SRPK1[69]. In this thesis, the combination of RNAi screening and high-content automatic imaging will be further exploited to identify the signaling landscape of TNBC metastatic modulators, but also systematically investigate the role of splicing factors in TNBC cell migration and proliferation.

Next to the introduction of the RNAi screening methods, the development of genomics and transcriptomics have led to an exponential increase in the identification of new biomarkers or potential disease targets. Public efforts such as The Cancer Genome Atlas (TCGA) provide amongst others transcriptomics data for 1097 breast tumors that resulted in improved classifications and discovery of subtype-specific biomarkers[70,71]. Furthermore, applying transcriptomics to cell lines in different culture conditions or after specific RNAi treatments enables the transcriptome-wide discovery of signaling pathways critical for the biological process of interest. In this thesis, we will provide a transcriptomics repository for 13 breast cancer cell lines cultured in both 2D and 3D conditions, but also perform transcriptomic analysis to uncover the downstream signaling pathways of critical migration determinants in TNBC cell lines. Moreover, we will use the patient-derived transcriptomics data provided by TCGA to link candidate expression to breast cancer progression and determine the co-regulation of the splicing machinery in primary breast tumors.

Aim and outline of the thesis

In this thesis, we aimed to better understand the underlying mechanisms involved in TNBC progression and metastasis formation and discover new targets to reduce breast cancer related deaths. In **chapter 2** we performed an imaging-based RNAi phenotypic cell migration screen in two highly motile TNBC cancer cell lines (Hs578T and MDA-MB-231) to provide a repository of signaling determinants that functionally drive TNBC cell motility. We screened 4,200 target genes and discovered 133 and 113 migratory modulators of Hs578T and MDA-MB-231, respectively. Interestingly, two modulators essential for cancer cell migration, amplified in

human breast tumors and associated with metastasis-free survival were proteins known to be involved in RNA splicing (PRPF4B and BUD31). This in combination with results from our previously performed RNAi migratory screen in lung cancer cells discovering splicing factor SRPK1 as a determinant for breast cancer metastasis formation[66], made us decide to focus on the role of RNA splicing in breast cancer progression and metastasis formation. In **chapter 3**, we summarized the current knowledge about splicing factors in breast cancer development and progression. In **chapter 4**, we performed a systemic analysis of the co-regulation of splicing factors using primary breast tumor RNA sequencing data. We identified co-regulated splicing factors that were associated with aggressive breast cancer phenotypes and metastasis formation that was not only restricted to breast cancer, increasing the global understanding of the role of the spliceosome in cancer development and progression. Next, the role of splicing factors in two major processes in cancer progression, cell migration and proliferation, was examined. In **chapter 5** we performed an imaging-based RNAi screen to systematically unravel the role of all 244 splicing factors in TNBC cell migration. A selective set of splicing factors was identified that upon knockdown specifically inhibited breast cancer cell migration without affecting proliferation. These splicing factors regulated specific alternative splicing events resulting in downregulation of critical migratory pathways. In **chapter 6** we performed an RNAi screen to unravel the role of all 244 splicing factors in breast cancer cell proliferation. We identified nine splicing factors that upon knockdown inhibited proliferation in two TNBC cell lines via a common mechanism: knockdown dysregulated sister chromatid cohesion, resulting in cells bearing double DNA content, stall in G1-S transition and subsequently cell death. In **chapter 7**, using RNA sequencing, we systematically compared the transcriptomes of 14 breast cancer cell lines cultured both in 2D and 3D conditions to unravel the reprogramming that is associated with the invasive phenotype of basal B TNBC in comparison to basal A and luminal breast cancer. Moreover, these cell line transcriptomes were compared to the human patient tumor transcriptome. Finally, we established a bioinformatics strategy to define the culture conditions that are most optimal to study the oncogenic pathway of interest and improve clinical translation of in vitro experiments. In **chapter 8**, I summarized all findings discussed in this thesis and provide an overview of the implications and future perspectives of these studies.