ENGLISH SUMMARY

Breast cancer is the most common and mortal cancer among females worldwide, which is recognized as a highly heterogeneous disease composed of several molecularly distinct subtypes. Chemotherapy is a crucial treatment for breast cancer, especially for the hormone-insensitive, advanced or metastatic breast cancer. Taxane-containing chemotherapy regimens has been shown with superiority compared to non-taxane regimens. Nowadays, taxanes, either alone or in combination have been commonly applied as first-line regimens for advanced breast cancer. However, the intrinsic or acquired resistance has limited the effectiveness of taxanes and led to recurrent and metastasis disease in breast cancer patients.

Many studies have been carried out to discover biomarkers for docetaxel resistance, with the aim to predict the response to docetaxel in breast cancer patients and thus guide the use of docetaxel-containing regimens or find solutions to eradicate the resistance. The overexpression of P-glycoprotein has been proposed to be a common mechanism of taxanes resistance. However, due to the inconsistency of clinical studies, there are still no predictive biomarkers for taxanes in clinical use.

Recently, the noncoding RNAs (including microRNA (miRNA), long non-coding RNA (IncRNA) and circular RNA (circRNAs)) have been shown to play important roles in chemotherapy resistance in cancers. In order to find more specific and applicable novel biomarkers for docetaxel resistance, it is imperative to fully profile the whole transcriptome in docetaxel-resistant breast cancer (short for DRBC).

In the former study carried out by our research group (Hansen et al., 2015), two DRBC cell lines (MDA-RES and MCF7-RES) have been obtained via exposure to increasing docetaxel concentrations. These cell lines have been proved to present resistance to docetaxel, and the overexpression and functional importance of P-glycoprotein was also validated in these two cell lines.

This PhD project utilizes these two former validated DRBC cell lines, which represent the two main breast cancer subtypes: the estrogen receptor (ER) positive and triple negative breast cancers. High through-put RNA sequencing technologies and
bioinformatic analyses were performed to detect the alterations in different classes of RNA molecules. This allowed us to decipher the relationships between different RNA molecules and identified novel putative biomarkers and mechanisms for docetaxel resistance for breast cancer.

**Manuscript I:** We performed ribosomal RNA-depleted RNA-sequencing on two DRBC cell lines (MDA-RES and MCF7-RES) and their docetaxel sensitive parental cell lines MDA-MB-231 and MCF-7. The expression of coding genes was analyzed and we identified 124 consistent significant up-regulated or down-regulated genes in both MDA-RES and MCF7-RES cells. The most studied gene in multidrug resistance, *ABCB1* (encodes P-glycoprotein), was dramatically up-regulated in both MDA-RES and MCF7-RES cells. However, we also identified other protein coding genes and pathways, which may contribute to the generation of docetaxel resistance. More important, we identified a group of significantly differentially expressed (SDE) lncRNAs, which were consistently up-regulated or down-regulated in both MDA-RES and MCF7-RES cells. The co-expression network and genomic location analyses pinpointed four overexpressed lncRNAs located within or near the *ABCB1* locus, which may up-regulate the expression of *ABCB1* via certain mechanisms. We also identified the lncRNA EPB41L4A-AS2 as a potential biomarker for docetaxel sensitivity. To the best of my knowledge, this is the first study to report on a global profile of mRNA and lncRNAs in DRBC cells. These findings have extended our knowledge about the mechanisms underlying *ABCB1* overexpression and docetaxel-resistance in breast cancer and provided potential predictive lncRNA biomarkers for docetaxel based treatments in breast cancer.

**Manuscript II:** To investigate the relationships between circRNAs and docetaxel resistance for breast cancer, we performed circRNAs analyses using the ribosomal RNA-depleted RNA-sequencing data from the Manuscript I. Our analyses identified different circRNA signatures between docetaxel-resistant and sensitive breast cancer cells and firstly discovered circRNAs generated by the multidrug resistant genes *ABCB1* and *EPHA3*. CircABCB1 was identified and validated as a dramatically up-regulated circRNA in both docetaxel resistant MCF7-RES and MDA-RES breast
cancer cells, whereas circEPHA3.1 and circEPHA3.2 were dramatically down-regulated in MCF7-RES cells. Furthermore, miRNA sequencing was performed, and the bioinformatics analysis showed that the circABCB1, circEPHA3.1 and circEPHA3.2 may probably soak the SDE miRNAs, which were associated with chemotherapy resistance and thus contribute to docetaxel resistance in breast cancer. Our data also indicated that PI3K-Akt and AGE-RAGE signaling pathway may probably mediate the circRNA-miRNA-mRNA regulation network. To the best of my knowledge, this represents the first established global profile of circRNAs in DRBC cells, which has shed some light into the roles of circRNAs in the docetaxel resistance of breast cancer.

To conclude, in this PhD project, we performed high through-put RNA and miRNA sequencing on two former validated DRBC cell lines and their docetaxel sensitive parental cell lines, respectively. A comprehensive bioinformatic analyses pipeline was set up to identify and quantify the coding and non-coding RNAs as well as decipher the interaction between different classes of RNA molecules. We identified a group of novel lncRNAs and circRNAs associated with the docetaxel resistant phenotype in these two docetaxel resistant cell lines, and uncovered potential regulators for the overexpression of P-glycoprotein. These novel potential biomarkers may be helpful to classify breast cancer patients with different response to docetaxel and find therapeutic targets to eradicate docetaxel resistance in breast cancer, and the imperative prerequisite is that their biological function and the clinical relevance should be thoroughly defined by both in vivo and in vitro studies.