Using insertional mutagenesis to identify breast cancer drivers and therapy resistance genes in mice = Insertie mutagenese voor het identificeren van genen betrokken bij de ontwikkeling van borsttumoren en therapie resistentie in muizen

Kas, S.M.

Appendices

English summary

Nederlandse samenvatting

Curriculum Vitae

List of publications

Acknowledgements

English summary

Breast cancer is the most common cancer diagnosed in women worldwide, with nearly 1.7 million new cases each year. The last decades it became clear that breast cancer is a complex and heterogeneous disease, which can be divided in multiple distinct subtypes that differ in terms of morphology, genetic alterations, molecular characteristics, responses to treatments and clinical outcomes. Despite the improved survival rates of human breast cancer patients due to the implementation of screening methods for early detection of tumors and the progress in treatment strategies, it remains difficult to successfully treat patients with tumors that progressed to a more advanced stage. To improve the clinical outcome of patients with invasive breast cancer it is crucial to identify the genes and pathways that drive tumor formation to increase our understanding of the disease and to find novel therapeutic vulnerabilities. In this thesis, we used different genetic approaches in mice to study the roles of candidate cancer genes and pathways in the development of invasive lobular breast carcinoma (ILC), which is the second most common breast cancer subtype. ILCs are hallmarked by the loss of E-cadherin and show overall favorable clinical features, including low histological grade, low proliferation rates and the majority of the tumors are hormone-receptor positive. However, ILCs also show typical invasive growth patterns in adjacent healthy tissue, which make them difficult to detect by palpation or mammography and patients are therefore likely to be diagnosed with a more advanced stage of the disease.

To illustrate the complexity of breast cancer heterogeneity we reviewed the current relevant literature in **chapter 1**, with the focus on the histopathological, genetic and molecular features of ILCs. Furthermore, we describe the advantages and limitations of genetically engineered mouse models to study the contribution of candidate cancer genes in tumor formation, progression and metastasis formation. We expect that recent advances in genetic engineering will significantly boost the validation of candidate cancer genes in mouse models that faithfully recapitulate human cancers. In line with this, the development of these next-generation mouse models will further improve our understanding in various aspects of tumor biology, including the crosstalk between cancer cells and the microenvironment and their contribution in response and resistance to novel anti-cancer drugs.

Large-scale sequencing studies are rapidly uncovering the mutational landscapes of human tumors. However, it remains challenging to discriminate between driver and passenger mutations for tumor formation. Putative cancer genes require validation in models that faithfully recapitulate human disease. In **chapter 2** we describe a novel strategy in mice to rapidly test the contribution of candidate cancer genes in the development of ILCs. In this strategy, we performed intraductal injections with high-titer lentiviruses to activate or inactivate genes that are involved in PI3K-AKT signaling, which is frequently activated in human ILCs. Using this approach, we were able to target ILC-initiating cells and demonstrated that combined inactivation of E-cadherin and activation of an oncogenic AKT variant resulted in the development of ILCs that reflected their human counterpart. Furthermore, we

also induced the formation of ILCs by CRISPR/Cas9-mediated somatic disruption of *Pten* in E-cadherin-deficient mammary epithelial cells in mice, highlighting the utility of this versatile platform to rapidly test putative tumor suppressor genes implicated in ILC.

Despite the fact that the majority of human ILCs are characterized by the loss of E-cadherin (encoded by *Cdh1*), mice with mammary-specific inactivation of *Cdh1* are not predisposed to develop mammary tumors, which suggests that mutations in additional genes are required to induce ILC formation. To identify these genes, we performed a *Sleeping Beauty* (*SB*)-mediated insertional mutagenesis screen in mice with mammary-specific inactivation of *Cdh1* (**chapter 3**). These mice showed multifocal mammary tumor formation of which the majority resembled human ILC based on morphology and gene expression. Analysis of common insertion sites identified several known and novel candidate cancer genes in the development of ILC, highlighting the strength of using insertional mutagenesis in mice for cancer gene discovery. Recurrent and mutually exclusive transposon insertions were identified in *Myh9* , *Ppp1r12a*, *Ppp1r12b* and *Trp53bp2* of which *MYH9*, *PPP1R12B* and *TP53BP2* were also frequently aberrated in human ILCs. Interestingly, these genes have been implicated in the regulation of the actin cytoskeleton, suggesting the identification of drivers of a novel oncogenic pathway involved in ILC formation.

In transposon-based insertional mutagenesis screens, the analysis of common insertion sites is typically performed by targeted DNA-sequencing of transposon insertions followed by heuristics to identify candidate cancer genes. However, this approach provides no direct evidence that transposon insertions truly affect their predicted targets. In **chapter 4** we developed IM-Fusion, an approach that uses single- and paired-end RNA-sequencing data to identify insertion sites from gene-transposon fusions. Using this approach, we were able to accurately identify transposon insertions and their predicted target genes in two separate transposon screens of mammary tumors and acute lymphoblastic leukemias, respectively. An advantage of IM-Fusion is that gene expression quantification can be combined with the insertion sites to determine the effect of a transposon insertion on its target gene, which can be used to prioritize the validation of candidate cancer genes. This approach provided valuable insights in the underlying mechanisms of PPP1R12A, PPP1R12B and TRP53BP2 to drive ILC formation (**chapter 3**), as strong clustering of transposon insertions in the respective genes suggested the expression of truncated transcripts. Analysis of gene expression before and after the insertion sites confirmed the overexpression of truncated *Ppp1r12a*, *Ppp1r12b* and *Trp53bp2* transcripts, which resulted in the expression of truncated proteins lacking various regulatory domains that could affect their function.

In **chapter 5** we present another example of the strength of using insertional mutagenesis screens to identify candidate cancer genes and to uncover additional biological functions of specific protein domains. In more than half of the *SB*-induced mouse ILCs we identified transposon insertions in *Fgfr2* (**chapter 3**), of which the majority was clustered in the intron directly upstream of the last exon resulting in the expression of C-terminal truncated *Fgfr2* transcripts. Using genetically

engineered mice, we show that mammary-specific expression of truncated FGFR2 variants induced rapid mammary tumor formation, whereas expression of full-length FGFR2 did not result in tumor development. Furthermore, we show that the C-terminus of FGFR2 harbors multiple domains that are essential in suppressing tumor formation, suggesting that several factors regulate the activity of FGFR signaling via the C-terminal tail.

In **chapter 6** we used insertional mutagenesis in mice to identify genes that confer resistance to FGFR-targeting therapy. We performed orthotopic transplantation of *SB*-induced mouse ILCs with active FGFR signaling and treated these tumors with the clinically relevant FGFR inhibitor AZD4547. These tumors regressed initially during treatment, but eventually developed resistance to the drug. By combining transcriptomic and insertion site analysis of AZD4547-resistant tumors, we identified several known and novel mechanisms of resistance, which all converged on reactivation of the canonical MAPK-ERK signaling pathway. Interestingly, two of these resistance mechanisms were only identified by acquired transposon insertions in AZD4547-resistant tumors, demonstrating that insertional mutagenesis in mice is a powerful tool to identify resistance mechanisms to anti-cancer targeting drugs.

Taken together, this thesis describes the identification of candidate cancer genes in ILC formation and genes that confer resistance to FGFR-targeting therapy by performing insertional mutagenesis screens in mice. Using these screens, we were also able to uncover biological processes underlying mammary tumor development, which highlights this powerful genetic approach for studying cancer biology in mouse models that recapitulate human disease.