Qiuqiong Tang
Dr. sc. hum.

DNA methylation array analysis identifies breast cancer associated methylation changes in peripheral blood DNA

Fach/Einrichtung: DKFZ
Doktormutter: Prof. Dr. rer. physiol. Barbara Burwinkel

Breast cancer (BC) is the most common malignant disease in women worldwide, and the leading cause of cancer mortality for women. Although a number of biomarkers have been developed to estimate the outcome in patients, the number of markers for early detection of BC and BC risk evaluation is till limited. Minimally invasive blood based biomarker BC early detection that can be measured repeatedly to provide disease information in a real time manner, are urgently needed.

In the present study, we first created an epigenome-wide DNA methylation profiling on peripheral blood DNA from 48 sporadic BC cases and 48 healthy controls with Infinium HumanMethylation 450K BeadChip, aiming to identify the strongest methylation changes in blood DNA that are associated with BC. The 450K data was analyzed in succession by two different methods. Firstly, the 450K data was analyzed with BMIQ normalization method, but the data was not adjusted with cell proportions. Secondly, the 450K data was re-analyzed with Functional normalization method, which was claimed to be with better performance than BMIQ. To test for the influence of cellular heterogeneity, the method developed by Houseman and the Reinius reference dataset were implemented to first estimate the proportions of six different sub cell type of leucocytes for each sample and then adjust for those in the beta regression.

Candidate CpG sites were selected by two strategies, which were based on the 450K data analyzed by the two different methods mentioned above, respectively. Although these two strategies with different cut-off points, the principle criteria within them were the p values for methylation differences between groups and the actual methylation difference between the compared groups. The identified CpG sites were first run for pre-test on a small sample cohort to verify the 450K data. Successfully verified CpG sites were considered for further validations. Three CpG sites selected by strategy I, which are located in *RPTOR*, *MGRN1* and *RAPSN* were further validated in three independent sample cohorts (Validation I: 102 controls vs 109 cases; Validation II: 189 controls vs 189 sporadic cases; Validation III: 250 controls vs 270 familial BC cases) by MassARRAY. One CpG site selected by strategy II, which is located in *PPP1R27* was further validated in two independent case-controls (Validation I: 151 controls vs 229 sporadic cases; Validation II: 379 controls vs 378 familial BC cases) study samples. These four CpG sites, as well as the adjacent CpG sites included in the regions analyzed by MassARRAY assay, were

successfully validated in different validation cohort samples. Results of all rounds of validation confirmed a significant decrease of methylation of these loci in peripheral blood DNA of both sporadic and familial BC cases compared to healthy controls. Decreased methylation intensities of investigated genes were associated with increased BC risk in our studied samples. Combination analysis was conducted to evaluate the value of our investigated methylation markers for distinguishing BC patients from healthy females and a four - gene DNA methylation panel was developed. The summary of the project was depicted in Figure 1.

Methylation data of our four - genes panel could differentiate the sporadic BC cases from controls, with an AUC of 0.79, and also could differentiating FBC patients from healthy females with an AUC of 0.69. Overall, our four - gene methylation panel could distinguish BC cases from healthy controls, with an AUC of 0.71, which appear to generate better predictions of risk than the well - known Gail model or GWAS SNPs alone, and if this can be verified. To sum up, we have identified four methylation markers (four - gene maker panel) for the first time in whole blood DNA. This would be of significant translational relevance since they may have the potential to be used in a clinical setting either as a prescreening assay to facilitate recommendations for further diagnostic tests or in combination with other established markers.

This thesis mainly focus on DNA methylation signatures identified from epigenome-wide DNA methylation profiling without prior knowledge of biological functions of investigated genes. Therefore, the present findings provided indications for further studies which may elucidate underlying biological mechanisms. Methylation changes of other genes, which did not show significant associations, may also have an impact on BC etiology and/or risk, but were disregarded due to stringent correction for multiple comparisons.

In summary, this thesis elucidated for the first time that decreased methylation of *RPTOR*, *MGRN1*, *RAPSN* and *PPP1R27* is associated with sporadic as well as familiar BC and might be associated with the etiology and/or risk of BC in women of European ancestry. These might be part of a non-invasive blood-based molecular marker set for the evaluation of BC risk or early detection. Further genetic and functional studies, especially prospective studies are still required to identify the causal relationships and the mechanisms underlying the observed associations.

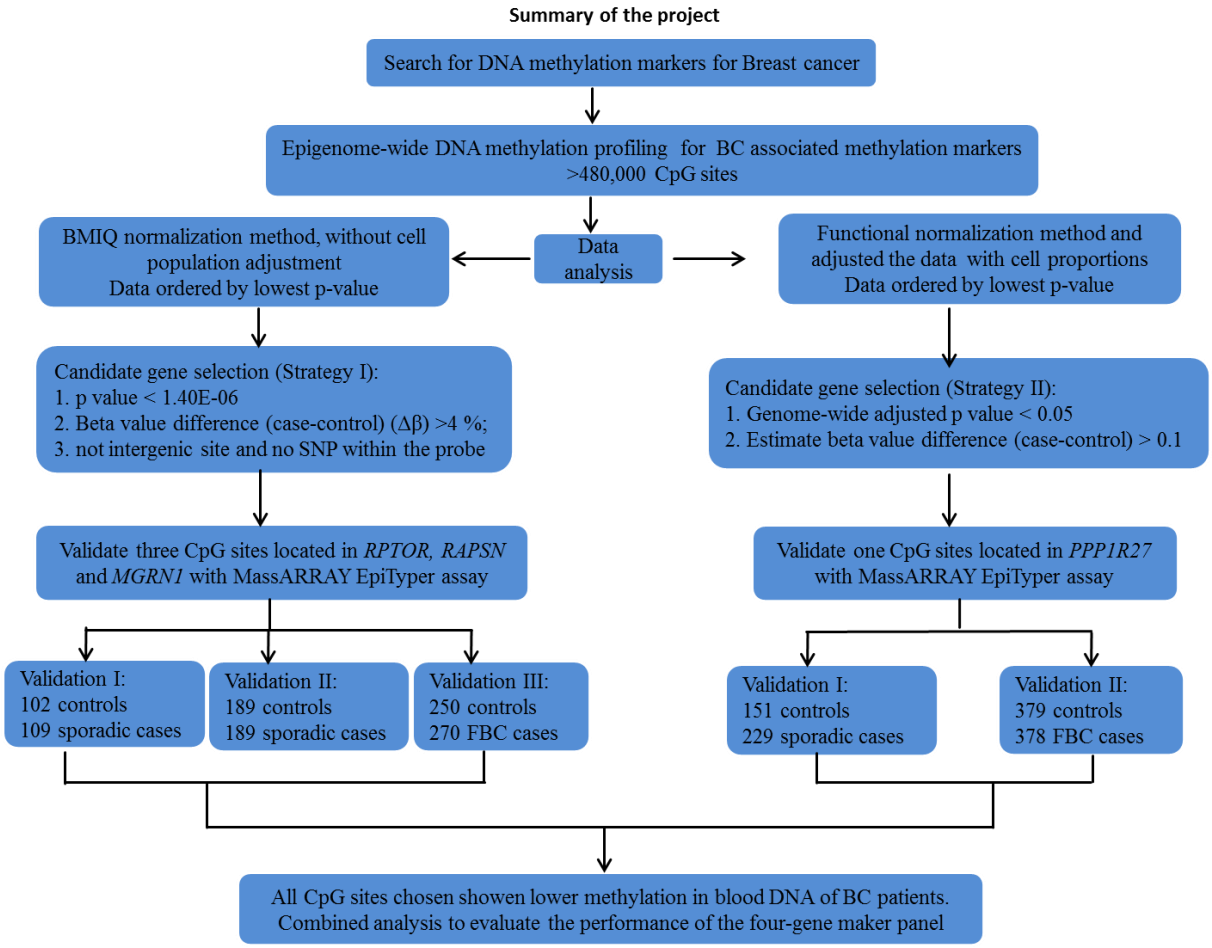

Figure 1. Summary of the project.