Xu Gao

Dr. sc. hum.

Smoking-related DNA methylation markers: relationship with aging-related health outcomes

Fach/Einrichtung: DKFZ (Deutsches Krebsforschungszentrum)

Doktorvater: Prof. Dr. med. Hermann Brenner

Tobacco smoking has been recognized as a modifiable risk factor for many aging-related health outcomes and a leading cause of preventable morbidity and mortality. Epigenetic research showed that smoking is one of the crucial environmental factors that could influence the DNA methylation, one of the main forms of epigenetic modification. With the advent of array-based technology, previous epigenome-wide association studies (EWASs) have discovered an increasing number of CpG sites that were not only associated with both current and lifetime smoking exposure quantitatively, but also with the adverse health effects of smoking.

Moreover, recent advances in aging research have also disclosed several aging-related parameters, including but not restricted to the DNA methylation age, telomere length (TL), frailty index (FI) and oxidative stress (OS), which could facilitate the disclosure of accelerated aging and its potential adverse health consequences. Regarding potential mechanisms between smoking and aging, a broadly accepted hypothesis was that smoking could produce a variety of oxidants and various toxic chemicals, which might be involved in accelerated aging. Nevertheless, underlying potential genetic or epigenetic mechanisms associated with smoking-related adverse health effects at the old age were only partly understood. Therefore, this dissertation aimed to provide better knowledge of the impact of smoking on aging-related health outcomes and potential underlying mechanisms; the primary objective was to investigate the associations of smoking and smoking-related DNA methylation patterns with aging-related health outcomes.

First, a systematic review was conducted to summarize the evidence on smoking-related DNA methylation changes in whole blood samples. Overall, 1460 smoking-related CpG sites were identified in 14 EWASs, of which 151 sites were detected in multiple (≥2) studies. These significant smoking-related genes were further assessed by specific methylation assays in three gene-specific methylation studies, and reflected not only current but also lifetime or long-term exposure to active smoking. This review suggested that DNA methylation signatures from whole blood samples may be the most powerful and highly informative biomarkers not only for current smoking but also for lifetime history of smoking.

Next, as lung cancer is the most common cancer caused by smoking and the most common form of cancer and cause of cancer-related death worldwide, this dissertation examined in the ESTHER study whether smoking could alter methylation of genes at lung cancer risk loci identified by genome-wide association studies. A total of 2854 corresponding CpG sites within 75 lung cancer related genomic

regions were selected for this analysis. Thirteen CpG sites located in 8 genes were successfully identified, current smoking was linked to a 0.74% to 2.4% decrease in DNA methylation compared to never smoking in 11 loci, and all but one showed significant associations (FDR <0.05) with lifetime cumulative smoking (pack-years). This study indicated that lung cancer susceptibility genes might be regulated by methylation changes in response to smoking exposure.

In a third step, the associations of smoking and smoking-related DNA methylation patterns with four aging-related parameters were assessed in the ESTHER study. A total of 66, 7, 9 and 71 of the 151 smoking-related CpG sites identified from the systematic review showed significant associations with DNA methylation age, TL, FI and OS, respectively and were used to construct corresponding smoking indices (SIs). Such SIs manifested monotonic associations with the changes of all aging-related parameters, which persisted and remained essentially unchanged even after adjusting for the smoking status in the regression models. These identified CpG sites could potentially be prognostic epigenetic markers of disproportional aging and aging-related health outcomes in response to smoking exposure.

In addition, this dissertation assessed how genetic variation might influence the smoking-related epigenetic changes. In this study, 70 out of 151 previously reported smoking-related CpG sites were found to be significantly associated with 192 SNPs within the 50kb search window of each locus. These SNPs were assigned as the methylation quantitative trait loci (mQTLs) of smoking-related CpG sites. The 192 mQTLs significantly influenced the active smoking-related DNA methylation changes, especially for the weakly/moderately smoking-related CpG sites, but were not directly associated with active smoking exposure or all-cause mortality. These results demonstrate that if not dealt with properly, the mQTLs might impair the power of epigenetic-based models of smoking exposure to a certain extent. This study further adds evidence for the complex interplays among genetic traits, epigenetic signatures and smoking exposure.

In summary, the investigations performed within the scope of this doctoral dissertation suggested that DNA methylation patterns are highly informative as potential mechanistic links between tobacco smoking and aging-related health outcomes and may therefore hold promises for disease risk assessment and prevention. Smoking could cause aging-related diseases by affecting the risk genes via DNA methylation and lead to the disproportionate aging through DNA methylation changes at specific CpG sites. As the susceptibility of the epigenome to smoking exposure strongly varies between individuals, understanding the inherited differences in epigenetic states of people reflected by the mQTLs might enhance the epigenetic-based assessments for smoking or smoking-related health outcomes by accounting for potential confounding from genetic background. Future studies should thus make big efforts towards bringing epigenome and genome data together to construct predictive models for the impact of smoking on accelerated aging. Overall, epigenetics is at the crossroads of genes and the environment. Further research based on even larger cohorts including longitudinal measurements of smoking exposure and epigenetic changes should be conducted to further clarify the biological and population-level relationships in this triangle. Such research should enable further enhanced smoking-related risk prediction and disclosure of smoking-related mechanisms in the pathogenesis of aging-related diseases.