Yan Zhang
Dr.sc.hum.

Smoking-associated DNA methylation biomarkers and their predictive values for smoking-related health risks

Promotionsfach: DKFZ
Doktorvater: Prof. Dr. med. Hermann Brenner

Tobacco smoking is a major cause of avoidable morbidity and premature mortality. Ascertainment of smoking exposure in epidemiological and clinical studies and practice mostly relies on self-report, which is prone to inaccuracy due to intentional underreporting and imperfect recall of lifetime exposure. Even though a number of biomarkers are well established, such as cotinine levels and DNA adducts, they exclusively reflect relatively short-term smoking exposure and are of limited use for quantifying lifetime cumulative exposure, which is the strongest determinant for smoking-related risks. Recent advances in research on DNA methylation may open new avenues in search for biomarkers reflecting long-term past smoking exposure which might have the potential to better quantify the adverse health effects of smoking, and to contribute to better understanding of the underlying mechanisms.

Recent epigenome-wide DNA methylation studies have consistently disclosed pronounced differences in blood DNA methylation of *F2RL3* (cg03636183), *AHRR* (cg21161138, cg05575921, cg23576855), *2q37.1* (cg21566642, cg01940273, cg05951221, cg06644428), and *6p21.33* (cg06126421) between smokers and nonsmokers of various ethnicities. However, these smoking-associated methylation biomarkers were only discovered very recently, and information on their associations with various smoking indicators and on their clinical implications had remained unclear. The aim of this dissertation was therefore to investigate the associations of methylation intensity at these CpG sites with both current and lifetime active smoking exposure, as well as with mortality outcomes and lung cancer risk.

In a large population-based prospective cohort study of older adults, *F2RL3* methylation was quantified in baseline blood samples of 5009 participants by MALDI-TOF mass spectrometry, and a genome-wide methylation screen, which covered all the assessed CpG sties, was conducted among 1000 participants by the Illumina 450K platform. Dose-response relationships of smoking exposure with methylation intensities were examined by restricted cubic spline regression. The associations of individual and aggregate methylation patterns and of smoking with mortality outcomes and with lung cancer incidence were assessed by Cox regression. Potential confounding factors were controlled for in all regression models.

Clear and consistent dose-response relationships with respect to current and lifetime smoking intensity as well as time since cessation were consistently observed for methylation at 6 CpGs [*F2RL3* (cg03636183), *AHRR* (cg05575921, cg21161138), *2q37.1* (cg21566642, cg01940273), and *6q21.22* (cg06126421)]: methylation intensity at each CpG site steeply decreased with increasing smoking intensity up to approximately 15 cigarettes per day and with increasing cumulative smoking up to approximately 30 – 40 pack-years, followed by a further gradual decrease at higher current and lifetime smoking intensity; among former smokers, methylation intensities steadily increased with time since cessation up to approximately 20 – 25 years after quitting and levelled off thereafter.

Hypomethylation of *F2RL3* was strongly associated with all-cause, cardiovascular, cancer, other and lung cancer mortality as well as lung cancer incidence, even after adjustment for smoking and multiple other covariates, showing adjusted hazard ratios (HR, 95% CI) of 2.60 (1.81 – 3.74), 2.45 (1.28 – 4.68), 2.94 (1.68 – 5.14), 2.39 (1.11 – 5.16), 7.13 (3.41 – 14.49), and 5.48 (2.60 – 11.55), respectively, for participants in the lowest quartile of methylation intensity compared to those in the highest quartile. All the observed associations with assessed outcomes were much stronger among men than among women. *F2RL3* methylation alone or in combination with pack-years was highly predictive for occurrence of lung cancer within 11-year follow-up, particularly in participants >65 years (optimism-corrected Harrell's C statistics: 0.83 and 0.86, respectively).

Demethylation of *AHRR, 6q21.22* and *2q37.1* was also associated with mortality outcomes. Methylation of the top 2 CpGs [*AHRR* (cg05575921 and *6q21.22* (cg06126421)] showed very strong associations with all-cause, cardiovascular and cancer mortality, with adjusted HRs of 3.59 (2.10 – 6.16), 7.41 (2.81 – 19.54), and 2.48 (1.01 – 6.08), respectively, for participants with methylation levels in the lowest quartile at both CpGs. Adding methylation at these 2 CpGs into a model that included the variables of the Systematic Coronary Risk Evaluation chart for fatal cardiovascular risk prediction raised Harrell's C statistics from 0.736 to 0.785, and reclassified a substantial proportion of individuals [net reclassification improvement (NRI), 14.08%; integrated discrimination improvement (IDI), 2.63%].

This dissertation demonstrates the utility of blood DNA methylation biomarkers in the study of smoking-related diseases. These novel methylation biomarkers are highly informative for both current and lifetime smoking exposure and strongly predictive for smoking-related health risks, and may therefore hold promises for disease risk assessment and disease prevention.