Genome-wide methylation profiling of biliary tract cancers reveals pathogenesis mechanisms and clinical applications

Ghada Nouairia[1], Martin Cornillet[1], Trine Folseraas[2], Hannes Jansson[3], Erik von Seth[1], Amanda Klein[1], Lina Lindström Älgå[1], Aiva Lundberg Båve[1], Christina Villard[1], Mette Vesterhus[2], Tom Hemming Karlsen[2], Ernesto Sparrelid[3], Annika Bergquist[1]

[1] Department of Medicine Huddinge, Unit of Gastroenterology and Rheumatology, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden.
[2] Norwegian PSC Research Centre, Department of Transplantation Medicine, Division of Surgery, Inflammatory Diseases and Transplantation, Oslo University Hospital Rikshospitalet, Oslo, Norway.
[3] Division of Surgery, Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden.

Background

Primary sclerosing cholangitis (PSC) is the most established risk factor for the development of biliary tract cancers; cholangiocarcinoma (CCA) and gallbladder cancer (GBC). These tumors are rare and often diagnosed at a late stage leading to dramatic survival rate. There are currently no efficient tools for the diagnosis, prognosis, and monitoring of CCA and GBC in PSC patients. The pathogenesis of both bile tract cancers and PSC are poorly understood, limiting the development of therapeutic tools. As epigenetic mechanisms, such as DNA methylation, are now recognized as a hallmark of cancer and constitute a dynamic process throughout the human lifetime, we hypothesize that disease-related DNA methylation alterations are detectable in the peripheral blood of patients. We aim to identify such alterations in the peripheral blood of the patients that (1) behold prognosis potential to predict severe outcomes of PSC and (2) detect CCA and GBC at an early-stage. The study is as well designed to unveil specific methylation signature of bi liary tract cancers and PSC to provide new biological insights.

AIM:

- Identify specific DNA methylations profile of CCA and GBC
- Evaluate diagnostic and prognostic potential of DNA methylation in bile tract cancers

Methods

DNA was purified from the whole blood of 110 individuals including cholangiocarcinoma (40), gallbladder cancer (21), PSC (29) with and without CCA and healthy donors (20). We used the largest currently available DNA methylation profiling array, Illumina Infinium EPIC beadchip on 850 000 sites across the genome, at the Bioinformatics and Expression Analysis (BEA) core facility at KI, Campus Flemingsberg. We also collected complete clinical data of the patients including the outcomes (overall survival for cancer patients, liver transplantation, BTC occurrence or death in PSC patients). The acquired DNA methylation data was analyzed to detect differentially methylated sites and regions of the genome related to each disease in comparison to control groups (either healthy donors or PSC). We used Mann-Whitney based statistical test and semi-supervised machine learning (ML) methods for variable selection (Random Forest based Recursive Feature Elimination and Elastic Net regression).

Table 1. Cross-sectional study design.
n: number of patients

Groups	Exploration	Validation
Cholangiocarcinoma (CCA)	39	82
Gallbladder Cancer (GBC)	20	51
Primary Sclerosing Cholangitis (PSC)	30	62
Healthy donors (HD)	20	-
Other conditions/disease control	-	54
External samples (CCA and PSC)	-	36

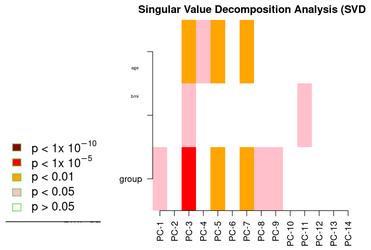

Figure 1. Overview of the main confounders (age and BMI) in the data. We checked several plausible confounders, known to affect DNA methylation, such as alcohol consumption, smoking, comorbities, etc. Only the factors with significant effects are visualized here. We adjusted the data for cell type composition and for age.

Results

We defined the DNA methylation profiles of GBC, CCA and PSC patients, including hypermethylated and hypomethylated regions of the DNA in each case. Using the differentially methylated positions of the DNA, we listed enriched genes and biological pathways that are significantly affected by DNA methylation, for each disease. We also identified novel genes, that are not known to be mutated in cancer, but that turned out to be associated to CCA and GBC through epigenetic changes, a phenomenon called epi-mutation.

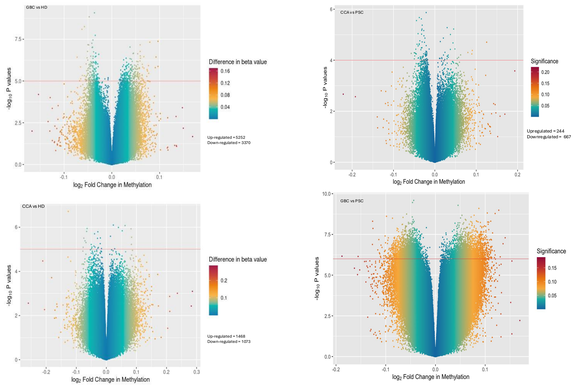

Figure 2. Volcano plots showing the proportion of differentially methylated sites when comparing bile tract cancers (CCA and GBC) to PSC or healthy individuals (HD). Results obtained by Mann-Whitney rank sum test.

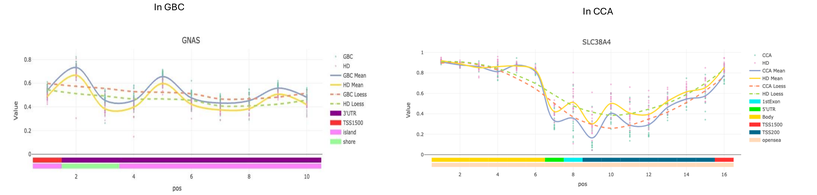

Figure 3. Example of genes enriched with differentially methylated sites in CCA (top) and GBC (bottom). SLC38A4: Related to hepatic cells that may exist in higher proportions in CCA peripheral blood, Known to be enriched in cancer (favorable) and is a prognostic marker in liver cancer (favorable), cancer enriched. GNAS is known to drive cancer through mutation in GBC.

We found around 2000 DNA methylation sites that either (1) have high statistical significance ($p < 0.01$), (2) occur in biologically significant genomic region, and/or (3) identified as diagnosis-associated sites by ML methods. Combinations of these sites can predict CCA and GBC (by comparison to healthy donors and to PSC patients) with high accuracy, measured by the receiver operating characteristic curve (ROC) based area under the curve (AUC). These combinations predict the prognosis of CCA in PSC patients and survival after surgical resection of the tumor. In addition, we identified metabolic pathways that are heavily impacted by DNA methylation in the studied diseases by performing enrichment analysis.

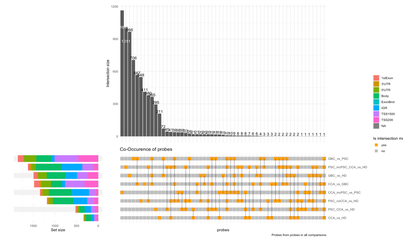

Figure 4. Probes that are specific for each comparison

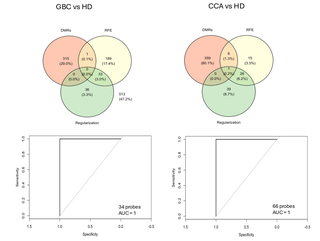

Figure 5. Few DNA methylation sites can be used to detect bile tract cancers

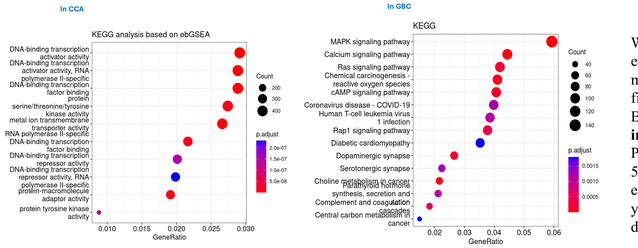

Figure 6. The most enriched pathways in differencially methylated regions (p < 0.05) in bile tract cancers.

We are following similar experimental and analytical methods to validate our findings on a larger cohort. Blood samples from **285 individuals** including 20 PSC patients' samples with 5 longitudinal sampling for each of them, for up to 10 years prior to CCA diagnosis.

Conclusion

We describe the first genome-scale DNA methylation profiles of biliary tract cancers and reveal significant differentially methylated sites and regions related to CCA and GBC as compared to PSC patients and to healthy individuals. Using machine learning methods, we identify several combinations of informative methylation sites with reliable diagnosis potential for CCA and GBC. We identified genes and pathways potentially involved in these biliary disorders, with some known to have a role in cancer and that might have implications for disease monitoring.

Ghada Nouairia
Eng., MSc., PhD

To get in contact:
Correspondence: ghada.nouairia@ki.se
0046731503125
Linkedin: Ghada Nouairia

Karolinska Institutet: https://ki.se/
Alfred Nobels Allé 8,
Stockholm, Sweden

Annika Bergquis
MD, PhD, Prof

Martin Cornillet
MSc., PhD

Visit Annika Bergquist group webpage: