Sarcoma classification by analyzing cell-free dna methylation profiles

Abstract Submitter: Florian Janke, Germany*

Co-Authors: Anja L. Riediger, Myesha Jahin, Simon Ogrodnik, Sabrina Gerhadt, Magdalena Görtz, Christoph E. Heilig, Daniel B. Lipka, Richard F. Schlenk, Stefan Fröhling, Holger Sültmann

*German Cancer Research Center & National Center for Tumor Diseases (NCT)

Abstract

Background: Sarcomas are rare and highly heterogeneous cancers arising from mesenchymal cells of soft tissue and bone. Owing to the presence of >80 histological subtypes and a scarcity of clearly defined molecular alterations, sarcoma classification poses a challenge for clinicians. Previous work demonstrated that tumor tissue DNA methylation profiling holds promise for improving sarcoma classification.

Objective: Evaluating the feasibility of sarcoma classification based on methylation profiles in cell-free DNA (cfDNA) in cases without available tumor tissue.

Methods: Plasma samples of 126 metastatic sarcoma patients (n=13 subtypes) and 36 non-cancer controls were collected. All samples were subjected to cell-free methylated DNA immunoprecipitation-sequencing (cfMeDIP-seq) and low-coverage whole genome sequencing to interrogate genome-wide DNA methylation patterns and copy number variations (CNVs), respectively. Additionally, publicly available Illumina 450k methylation array data from sarcoma tissues (n=1,477) were analyzed to identify subtype-specific methylation patterns and develop a sarcoma classifier.

Results: For 8 of 13 subtypes, both tissue (n=477) and cfDNA (n=92) methylation data were available. First, we identified subtype-specific DNA methylation signals by selecting the top 100 differentially methylated regions (DMRs) for each pairwise comparison across the 8 classes (1,786 DMRs). Next, we created 17,172 synthetic patients by mixing the cfMeDIP-seq data of every healthy donor cfDNA sample once with every sarcoma tissue dataset (at varying tumor DNA fractions). Then a gradient-boosted tree classifier was built based on 80% of the synthetic patients. The classifier performed accurately on the 20% held-out in silico mixtures (multi-class AUC = 0.945). When applied to cfDNA, no prediction was possible (‘unclassified’) for 38/92 of samples, and 26/92 sarcomas were misclassified. 59 of the 64 mis-/unclassified samples had no or few detectable CNVs (copy number abnormality [CPA] score <2) suggesting low circulating tumor DNA (ctDNA) content. However, after stratifying for samples with high ctDNA content (CPA ≥2), 23/28 sarcomas were correctly classified.

Conclusion: This project highlights the feasibility of cfDNA-based sarcoma classification from DNA methylation data, especially in patients with high ctDNA content.

Do you have any conflicts of interest?

No, I do not have a conflict of interest.