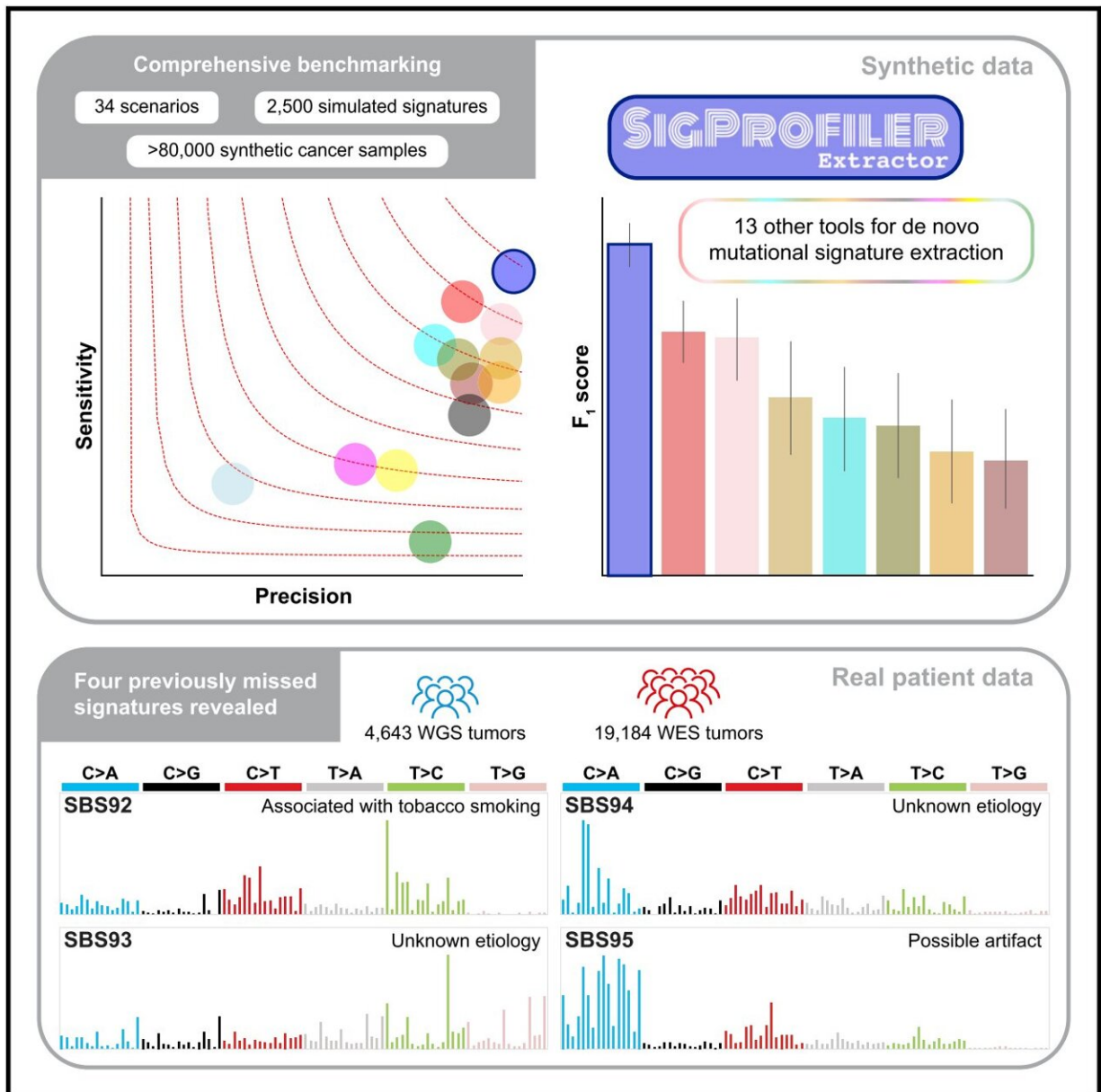


# Mutational signature linking bladder cancer and tobacco smoking found with new AI tool

September 26 2022



Graphical abstract. Credit: *Cell Genomics* (2022). DOI: 10.1016/j.xgen.2022.100179

Researchers at the University of California San Diego have for the first time discovered a pattern of DNA mutations that links bladder cancer to tobacco smoking. The discovery was made possible thanks to a powerful new machine learning tool that the team developed to find patterns of mutations caused by carcinogens and other DNA-altering processes.

The work, published Sept. 23 in *Cell Genomics*, could help researchers identify what [environmental factors](#), such as exposure to tobacco smoke and UV radiation, cause cancer in certain patients.

Each of these environmental exposures alters DNA in a unique way, generating a specific pattern of mutations, called a mutational signature. If a signature is found in the DNA of a patient's [cancer cells](#), the cancer can be traced back to the exposure that created that signature. Knowing which mutational signatures are present could also lead to more customized treatments for a patient's specific cancer.

In this study, researchers found a mutational signature in the DNA of [bladder cancer](#) that is linked to tobacco smoking. The finding is significant because a mutational signature from tobacco smoking has been detected in lung cancer, but not yet in bladder cancer.

"There is strong epidemiological evidence tying bladder cancer to tobacco smoking. We even see a specific mutational signature in other tissues—such as the mouth, esophagus and lungs—that are directly exposed to tobacco carcinogens," said study senior author Ludmil Alexandrov, professor of bioengineering and cellular and [molecular medicine](#) at UC San Diego. "The fact that we weren't finding this

signature in the bladder was strange."

Alexandrov and colleagues now show that there is a mutational signature from tobacco smoking in bladder cancer, and it's different to the signature found in [lung cancer](#). Moreover, they show that this signature is also found in normal bladder tissues of tobacco smokers who have not developed bladder cancer. The signature was not found in the bladder tissues of non-smokers.

"What this signature tells us is that certain mutations in your DNA are due to exposure to [tobacco smoke](#)," said study co-first author Marcos Diaz-Gay, a postdoctoral researcher in Alexandrov's lab. "It doesn't necessarily mean that you have cancer. But the more you smoke, the more mutations accumulate in your cells, and the more you increase your risk for developing cancer."

## **Made possible by next-generation machine learning**

The researchers found the tobacco signature with a next-generation machine learning tool developed by Alexandrov's lab. The team says it is the most advanced, automated bioinformatics tool for extracting mutational signatures directly from large amounts of genetic data.

"This is a powerful machine learning approach to recognize patterns of mutations and separate them from genomic data," said Alexandrov. "It takes those patterns and deciphers them, so that we can see what the mutational signatures are and match them with their meaning."

He compared the machine learning approach to picking out individual conversations at cocktail party.

"You have multiple groups of people talking all around you, and you are only interested in hearing certain individuals speaking," he said. "Our

tool essentially helps you do that, but with cancer genetic data. You have multiple people around the world exposed to different environmental mutagens, and some of those exposures are leaving imprints on their genomes. This tool goes through all that data to pick out what are the processes that cause the mutations."

The tool was used to analyze 23,827 sequenced human cancers. It found four mutational signatures—including the one in bladder cancer tied to [tobacco smoking](#)—that had not been detected by any other tool. The three other signatures, found in stomach, colon and liver cancers, still warrant further study to see what processes caused them.

To show how powerful their tool is, the researchers put it to the test against 13 existing bioinformatics tools. The tools were assessed for their ability to extract mutational signatures from more than 80,000 synthetic cancer samples. The tool that Alexandrov's team developed had outperformed all the others. It detected 20 to 50% more true positive signatures, with five times less false positive signatures. It even performed well when analyzing noisy data, whereas the other tools failed.

"In bioinformatics, this is the first time that such a comprehensive benchmarking has been done on this scale for mutational signature extraction," said Diaz-Gay. "It is a huge undertaking, comparing many tools across many datasets."

## **Creating a more user-friendly and personalized tool**

The team's ultimate goal is to create a web-based tool that more researchers can use and as a result, profile more patients.

"Right now, this tool requires bioinformatics expertise to run it," said Alexandrov. "What we want is to create a user-friendly version on the

web, where researchers can just drop in a patient's mutations, and it immediately gives you the set of [mutational signatures](#) and what processes caused them."

"Our idea for the future is to leverage this tool to analyze patients on an individual level," said Diaz-Gay.

**More information:** S.M. Ashiqul Islam et al, Uncovering novel mutational signatures by de novo extraction with SigProfilerExtractor, *Cell Genomics* (2022). [DOI: 10.1016/j.xgen.2022.100179](https://doi.org/10.1016/j.xgen.2022.100179)

Provided by University of California - San Diego

Citation: Mutational signature linking bladder cancer and tobacco smoking found with new AI tool (2022, September 26) retrieved 27 April 2024 from <https://medicalxpress.com/news/2022-09-mutational-signature-linking-bladder-cancer.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.