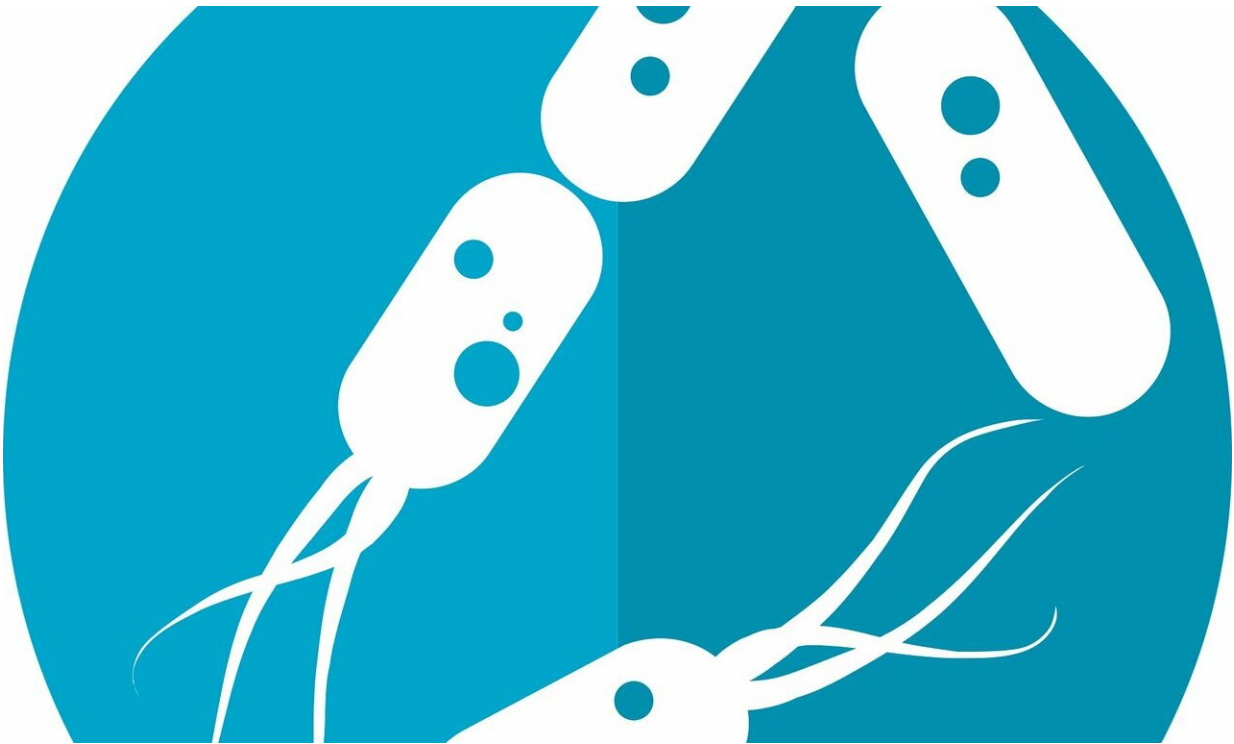


# New computer program aims to reduce DNA contamination in microbial samples

February 6 2019, by Tracey Peake

---



Credit: CC0 Public Domain

DNA sequencing of microbial samples can give researchers and medical professionals a wealth of information about microbiomes – the communities of microorganisms that inhabit our bodies and the environments all around us. Understanding the microbiome can aid our understanding of what ails us and why. But what happens when

microbial samples are contaminated with DNA from other sources?

"Contamination in [microbiome](#) sequencing can lead to false findings," says Ben Callahan, assistant professor of microbiomes and microbial communities at NC State. "For instance, researchers recently thought they had discovered several new microbes that could predict pre-term birth, but when they dug deeper those microbes turned out to be [contaminants](#). So these errors aren't inconsequential."

Getting samples that are completely free of contaminants from their surroundings isn't possible – researchers expect some level of contamination in most sequenced microbiome samples. However, contamination has a much larger impact when working with lower biomass samples (e.g. airway samples) than when working with high biomass samples (e.g. fecal samples), because in high biomass samples the legitimate microbial population overwhelms contamination.

Callahan and colleagues have created an open-source software package, called Decontam, which identifies contaminants in a [sample](#) using statistical patterns of the frequency and presence of contaminants versus non-contaminants. Contaminants appear at higher frequencies in low-concentration samples and often appear in negative controls.

Current methods for controlling contamination – including specialized lab practices or eliminating any unusual or rare microbial species from the sample – have significant drawbacks, as they can be costly, time-intensive, and could eliminate legitimate microbes from the sample. Additionally, these methods do not completely remove contaminants.

"Our method is simple, fast and cost-effective," Callahan says. "It is an algorithm that uses a simple binary classifier to distinguish between contaminants and noncontaminants on the basis of two patterns across samples: contaminants will increase in frequency as the amount of input

DNA decreases, and contaminants will be present in a higher fraction of negative control samples.

"Also, our method requires no additional data beyond what is typically generated in microbiome-sequencing experiments."

In testing, Decontam reduced the number of sequencing reads derived from contaminants by upwards of 99 percent in data collected from human mouths. The [method](#) was particularly effective at identifying and removing abundant [contaminants](#) that are the most likely to interfere with subsequent analysis.

You can find Decontam here: [github.com/benjineb/decontam](https://github.com/benjineb/decontam)

Callahan's work also appears in the journal *Microbiome*.

**More information:** Nicole M. Davis et al. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data, *Microbiome* (2018). [DOI: 10.1186/s40168-018-0605-2](#)

Provided by North Carolina State University

Citation: New computer program aims to reduce DNA contamination in microbial samples (2019, February 6) retrieved 26 April 2024 from <https://phys.org/news/2019-02-aims-dna-contamination-microbial-samples.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.