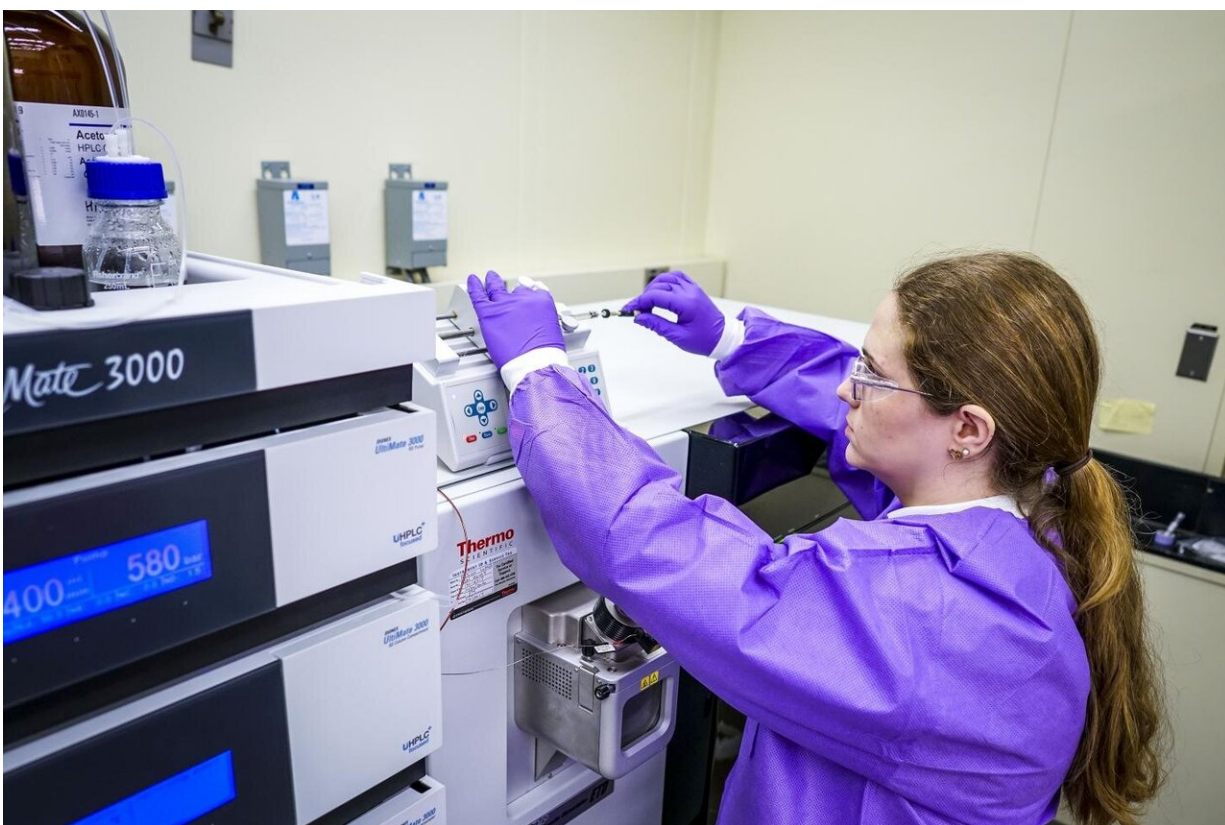


# New 'fingerprints' added to chemical identification database

June 17 2020

---



NIST research chemist Kelly Telu injects a sample into a mass spectrometer, a laboratory instrument that scientists use to identify unknown chemical compounds. Credit: M. Delorme/NIST

The National Institute of Standards and Technology (NIST) has updated

its database of chemical fingerprints, called mass spectra, that are used to identify unknown chemical compounds. The NIST Mass Spectral Library and its new version, called NIST20, is used in health care, drug discovery, foods and fragrances, oil and natural gas, environmental protection, forensic science and almost every other industry that manufactures or measures physical stuff.

"If you have a mysterious substance—you have no idea what it is—you generate its fingerprints then run those prints through our library," said NIST biostatistician Tytus Mak. "If you find a match, you know what the substance is."

Those chemical fingerprints are generated using a laboratory instrument called a [mass spectrometer](#) that breaks molecules into pieces then lines those pieces up on a graph according to their mass. The resulting mass spectrum appears as a series of vertical lines that form a unique pattern for each compound.

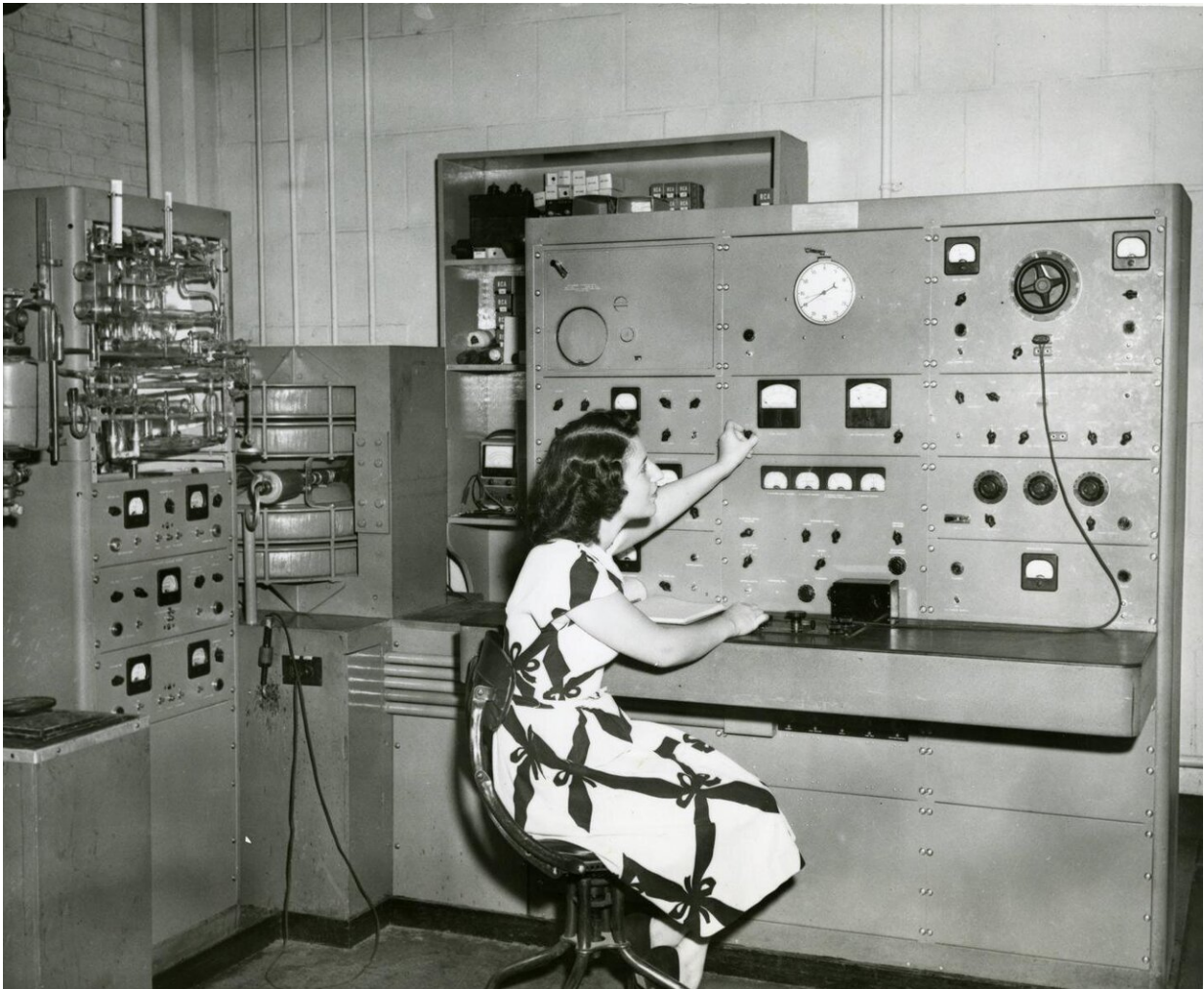
The NIST Mass Spectral Library comes pre-installed on many instruments, and users can purchase the update from their instrument manufacturer or other distributors. Collections of mass spectra used in specialized areas of research can be downloaded for free from the NIST website.

Mass spectrometry is particularly useful for identifying organic compounds—the building blocks of life. Part of Mak's role in this project was to decide, of the countless organic compounds out there, which ones to include in the library.

To do this, he scoured the catalogs of chemical manufacturers and lists of important compounds published by private companies, government agencies and scientific researchers. He then prioritized the compounds based on their relative importance and the cost of purchasing samples for

analysis.

This update includes more than 14,000 human and plant metabolites. Those are the substances formed when living things break down food, drugs or their own tissue, such as when you burn fat by exercising. Medical tests often involve identifying metabolites in blood or urine. Plant metabolites make up an even larger universe of chemical compounds. They are in everything we eat and are important in the agricultural sector.



In this 1948 photo, a NIST staff member operates an early mass spectrometer.

Credit: NIST

The update also included pesticides and environmental contaminants, chemicals used in manufacturing such as lubricants and surfactants, pharmaceutical drugs and illicit drugs such as new varieties of fentanyl, the drug that is driving a nationwide overdose epidemic.

After NIST purchased samples of the compounds, chemists ran them through carefully calibrated [mass](#) spectrometers. They did this on different instruments under varying conditions, producing multiple [mass spectra](#) for each compound. In keeping with NIST's high standards as the nation's measurement lab, a team of experts then analyzed the data to ensure high accuracy and precision.

"We carefully acquire and curate the data so users can have high confidence in their identifications," said NIST computational biologist Sara Yang, who worked on quality control.

The NIST Mass Spectral Library, which is among the larger commercially available libraries and is widely used, has two main components. The Electron Ionization (EI) Library is used for identifying volatile compounds such as those you can smell in air.

Roughly 40,000 new compounds have been added to this library, for a total of over 300,000. The Tandem Library is used to identify heavier compounds in liquids such as groundwater or blood. This library has almost doubled in size to more than 30,000 compounds and includes 1.3 million spectra.

Organic compounds are like Tinkertoys made mostly of carbon, hydrogen, oxygen and nitrogen atoms. They can be put together in an

endless number of ways. The diversity of life on Earth exists because of the vast possibilities of organic chemistry. Of all the [organic compounds](#) known and unknown, the NIST library has barely scratched the surface.

And the number of important compounds will continue to grow. There will always be new species of microbes to discover that might cause a new disease or produce a life-saving drug. And scientists will continue synthesizing new compounds, from [chemical](#) weapons to cures for cancer.

The job of updating the NIST Mass Spectral Library with new [compounds](#) will continue. But for now, chemists can easily identify tens of thousands more of them.

*This story is republished courtesy of NIST. Read the original story [here](#).*

Provided by National Institute of Standards and Technology

Citation: New 'fingerprints' added to chemical identification database (2020, June 17) retrieved 27 June 2024 from

<https://phys.org/news/2020-06-fingerprints-added-chemical-identification-database.html>

<p>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.</p>
--