

From 1 to 1,000s: Solving the Mysteries of Coronavirus With Genetic Fingerprints

Scientists traced the virus brought to the Seattle area in January. They were astonished to learn that the same branch of the virus traveled on through at least a dozen states and to other parts of the world.

By [Mike Baker](#) and [Sheri Fink](#)

April 22, 2020, 3:00 a.m. ET

SEATTLE — As the coronavirus outbreak consumed the city of Wuhan in China, new cases of the virus began to spread out like sparks flung from a fire.

Some landed thousands of miles away. By the middle of January, one had popped up in Chicago, another one near Phoenix. Two others came down in the Los Angeles area. Thanks to a little luck and a lot of containment, those flashes of the virus appear to have been snuffed out before they had a chance to take hold.

But on Jan. 15, at the international airport south of Seattle, a 35-year-old man returned from a visit to his family in the Wuhan region. He grabbed his luggage and booked a ride-share to his home north of the city.

The next day, as he went back to his tech job east of Seattle, he felt the first signs of a cough — not a bad one, not enough to send him home. He attended a group lunch with colleagues that week at a seafood restaurant near his office. As his symptoms got worse, he went grocery shopping near his home.

Days later, after the man became the first person in the United States to test positive for the coronavirus, teams from federal, state and local agencies descended to contain the case. Sixty-eight people — the ride-share driver at the airport, the lunchmates at the seafood restaurant, the other patients at the clinic where the man was first seen — were monitored for weeks. To everyone's relief, none ever tested positive for the virus.

But if the story ended there, the arc of the coronavirus's sweep through the United States would look much different.

As it turned out, the genetic building block of the virus detected in the man who had been to Wuhan would become a crucial clue for scientists who were trying to understand how the pathogen gained its first, crucial foothold.

Working out of laboratories along Seattle's Lake Union, researchers from the University of Washington and the Fred Hutchinson Cancer Research Center rushed to identify the RNA sequence of the cases from Washington State and around the country, comparing them with data coming in from around the world.

Using advanced technology that allows them to rapidly identify the tiny mutations that the virus makes in its virulent path through human hosts, the scientists working in Washington and several other states made two disconcerting discoveries.

The first was that the virus brought in by the man from Wuhan — or perhaps, as new data has suggested, by someone else who arrived carrying a nearly identical strain — had managed to settle into the population undetected.

Then they began to realize how far it had spread. A small outbreak that had established itself somewhere north of Seattle, they realized as they added new cases to their database, was now responsible for all known cases of community transmission they examined in Washington State in the month of February.

And it had jumped.

A genetically similar version of the virus — directly linked to that first case in Washington — was identified across 14 other states, as far away as Connecticut and Maryland. It settled in other parts of the world, in Australia, Mexico, Iceland, Canada, the United Kingdom and Uruguay. It landed in the Pacific, on the Grand Princess cruise ship.

The unique signature of the virus that reached America's shores in Seattle now accounts for a quarter of all U.S. cases made public by genomic sequencers in the United States.

With no widespread testing available, the high-tech detective work of the researchers in Seattle and their partners elsewhere would open the first clear window into how and where the virus was spreading — and how difficult it would be to contain.

Latest Updates: Coronavirus Outbreak in the U.S.

- [First coronavirus death in U.S. was weeks earlier than thought; warning of a 'second wave' this winter.](#)
- [Trump backs away from a broad ban on immigration, but many will still feel effect of restrictions.](#)
- [Disease detectives learn more about the virus's spread — and why it will be so hard to contain.](#)

[See more updates](#)

Updated 3m ago

Even as the path of the Washington State version of the virus was coursing eastward, new sparks from other strains were landing in New York, in the Midwest and in the South. And then they all began to intermingle.

A jigsaw puzzle

The researchers in Seattle included some of the world's most renowned experts on genomic sequencing, the process of analyzing the letters of a virus's genetic code to track its mutations. Before the outbreak, one of the labs had done more sequencing of human coronaviruses than anywhere else in the world — 58 of them.

When a virus takes hold in a person, it can replicate billions of times, some of those with tiny mutations, each new version competing for supremacy. Over the span of a month, scientists have learned, the version of the novel coronavirus moving through a community will mutate about twice — each one a one-letter change in an RNA strand of 29,903 nucleotides.

The alterations provide each new form of the virus with a small but distinctive variation to its predecessor, like a recipe passed down through a family. The mutations are so small, however, that it is unlikely that one version of the virus would affect patients differently than another one.

The virus originated with one pattern in Wuhan; by the time it reached Germany, three positions in the RNA strand had changed. Early cases in Italy had two entirely different variations.

For each case, the Seattle researchers compile millions of fragments of the genome into a complete strand that can help identify it based on whatever tiny mutations it has undergone.

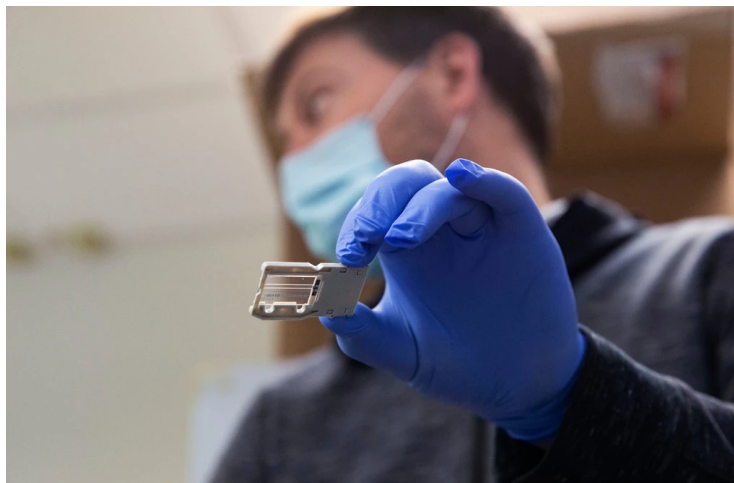
“What we're essentially doing is reading these small fragments of viral material and trying to jigsaw puzzle the genome together,” said Pavitra Roychoudhury, a researcher for the two institutions working on the sequencing in Seattle.

With some viruses, the puzzles are more challenging to assemble. The virus that causes Covid-19, she said, “was relatively well behaved.”

Researchers looked closely at the man who had flown in from Wuhan, who has not been publicly identified and did not respond to a request to speak to The New York Times.

They confirmed he had brought a strain of the virus that was already extending broad tentacles — from the Wuhan area to Guangdong on China's Pacific coast to Yunnan in the mountainous west. Along the way, its signature varied significantly from the version of the virus that spread in Europe and elsewhere: Its mutations were at positions 8,782, 18,060 and 28,144 on the RNA strand.

That gave Dr. Roychoudhury and the scientists around the country she has been working with the unique ability to see what the contact tracers in Seattle had been unable to: the invisible footprints of the pathogen as it moved.



A flow cell used for sequencing the coronavirus at a lab in Seattle. Ruth Fremson/The New York Times

An alarming find

On the hunt for the virus's path through the United States, one of the first signposts came on Feb. 24, when a teenager came into a clinic with what looked like the flu. The clinic was in Snohomish County, where the man who had traveled to China lived. Doctors gave the teenager a nasal swab as part of a tracking study that was already being done on influenza in the region.

Sign up to receive an email when we publish a new story about the **coronavirus outbreak**.

Sign Up

Only later did they learn that the teenager had not had the flu, but the coronavirus. After the diagnosis, researchers in Seattle ran the sample through a sequencing machine. Trevor Bedford, a scientist at the Fred Hutchinson Cancer Research Institute who studies the spread and evolution of viruses, said he and a colleague sipped on beers as they waited for the results to emerge on a laptop.

It confirmed what they had feared: The case was consistent with being a direct descendant of the first U.S. case, from Wuhan.

The teenager had not been in contact with the man who had traveled to Wuhan, so far as anyone knew. He had fallen ill long after that man was no longer contagious.

Additional sequencing in the days afterward helped confirm that other cases emerging were all part of the same group. This could only mean one thing: The virus had not been contained to the traveler from Wuhan and had been spreading for weeks. Either he had somehow spread it to others, or someone else had brought in a genetically identical version of the virus.

That latter possibility has become more likely in recent days, after new cases entered into the researchers' database showed an interesting pattern. A virus with a fingerprint nearly identical to the Wuhan traveler's had shown up in cases in British Columbia, just across the border from Washington State, suggesting to Dr. Bedford that it might not have been the first Wuhan traveler who had unleashed the outbreak.

Either way, the number of cases emerging around the time the teenager's illness was identified indicated that the virus had been circulating for weeks.

'I probably exposed a lot of people'

On its path through Washington State, one of the virus's early stops appears to have been at a square dance on Feb. 16 in the city of Lynnwood, midway between Seattle and Everett.

It was a full month since the Wuhan traveler's arrival. A couple dozen square dancers had gathered for a pie and ice cream social, capping off a series of practices and events from all over the region over the course of a three-day weekend.

Three groups of square dancers swung through promenades and allemandes — huffing and sweating to “Free Ride” and “Bad Case of Loving You.”

Stephen Cole, who was the dance caller that night, said he did not recall anyone showing signs of illness. But over the next few days, he and a woman who had been cuing the dance fell ill.

Another dancer, Suzanne Jones, had attended a class with Mr. Cole the day before. By the next weekend, Ms. Jones said, she started to feel symptoms she dismissed as allergies, since she had noticed the scotch broom starting to bloom.

After resting for a couple of days, Ms. Jones felt better and drove from her home in Skagit County more than 100 miles south to visit her mother in Enumclaw, helping pack some belongings for storage. On the way back, she visited the strip malls in Renton, then a store in Everett, then a laundromat in Arlington. She stopped to apply for a job with the Census Bureau.

“I probably exposed a lot of people that day,” she said.

Ms. Jones only realized it could be something more than allergies after getting a notification on March 2 that one of her square-dancing friends had died of the coronavirus as the outbreak began to emerge. She too tested positive.

There was minimal coronavirus testing in the United States during February, leaving researchers largely blind to the specific locations and mutations of the spread that month. The man who had traveled from Wuhan was not at the dance, nor was anyone else known to have traveled into the country with the coronavirus. But researchers learned that the virus by then was already spreading well beyond its point of origin — and all the cases of community transmission that month were part of that same genetic branch.

There was another spreading event. On the Saturday after the dance, a group of friends packed the living room of a one-bedroom apartment in Seattle, sharing homemade food and tropical-themed drinks.

Over the following days, several people began coming down with coronavirus symptoms. “Among people who attended, four out of every 10 got sick,” said Hanna Oltean, an epidemiologist with the Washington State Department of Health.

Several people passed on the virus to others. By late March, the state health department had documented at least three generations of “transmission occurring before anyone was symptomatic,” Ms. Oltean said.

By then, it was becoming clear that there were probably hundreds of cases already linked to the first point of infection that had been spreading undetected. It left a lingering question: If the virus had this much of a head start, how far had it gone?



The large outbreak on the Grand Princess, a researcher said, could probably be traced to a single person linked to the Washington State cluster. Jim Wilson/The New York Times

Spreading beyond Seattle

As cases of the virus spread, scientists in other states were sequencing as many as they could. In a lab at the University of California, San Francisco, Dr. Charles Chiu looked at a range of cases in the Bay Area, including nine passengers from the Grand Princess cruise ship, which had recently returned from a pair of ill-fated sailings to Mexico and Hawaii that left dozens of passengers infected with the coronavirus.

Dr. Chiu was stunned by his results: Five cases in the San Francisco area whose origins were unknown were linked back to the Washington State cluster. And all nine of the Grand Princess cases had a similar genetic link, with the same trademark mutations — plus a few new ones. The massive outbreak on the ship, Dr. Chiu believed, could probably be traced to a single person who had developed an infection linked to the Washington State cluster.

But it did not stop with the Grand Princess. David Shaffer, who had been on the first leg of the cruise with members of his family, said passengers on that leg did not discover until after they disembarked that the coronavirus had been aboard — when they learned that a fellow passenger had died.

He and his family felt fine when they returned to their home in Sacramento, he said, and when he started feeling sick the next day, on Feb. 22, he at first assumed it was a sinus infection.

Days later, he was tested and learned he had the coronavirus. His wife later tested positive, too, as did one of his sons and one of his grandsons, who had not been on the cruise.

Dr. Chiu remembers going over the implications in his head. “If it’s in California and it’s in Washington State, it’s very likely in other states.”

U.S. flare-ups abound

The same day Mr. Shaffer got sick, another person landed at Raleigh-Durham International Airport in North Carolina, having just visited the Life Care Center nursing home in Kirkland, which would become a center of infection. At the time, there were growing signs of a respiratory illness at the facility, but no indication of the coronavirus.

A few days later, the traveler began feeling ill, but with no sign that it might be anything serious, he went out for dinner at a restaurant in Raleigh. Just then, officials in Washington State began to report a coronavirus outbreak at Life Care Center. The person in North Carolina tested positive a few days later — the first case in the state.

By the middle of March, a team at Yale gathered nine coronavirus samples from the Connecticut region and put them through a portable sequencing machine. Seven came back with connections to Washington State.

“I was pretty surprised,” said Joseph Fauver, one of the researchers at the lab. At the time, he said, it suggested that the virus had been spreading more than people had initially believed.

In sequencing more recent cases, the researchers have found cases emanating from a larger cluster, with its own distinct genetic signature, originating in the New York area.

A group of cases throughout the Midwest, first surfacing in early March, appears to have roots in Europe. A group of cases in the South, which emerged around the same time, on March 3, appears like a more direct descendant from China.

But of all the branches that researchers have found, the strain from Washington State remains the earliest and one of the most potent.

It has surfaced in Arizona, California, Connecticut, the District of Columbia, Florida, Illinois, Michigan, Minnesota, New York, North Carolina, Oregon, Utah, Virginia, Wisconsin and Wyoming, and in six countries.

And new cases are still surfacing.

One of the enduring mysteries has been just how the virus managed to gain its first, fatal foothold in Washington.

Did the contact tracers who followed the steps of the man who had traveled from Wuhan miss something? Did he expose someone at the grocery store, or touch a door handle when he went to the restaurant near his office?

In recent days, the sequencing of new cases has revealed a surprising new possibility. A series of cases in British Columbia carried a genetic footprint very similar to the case of the Wuhan traveler. That opened up the possibility that someone could have carried that same branch of the virus from Wuhan to British Columbia or somewhere else in the region at nearly the same time. Perhaps it was that person whose illness had sparked the fateful outbreak.

But who? And how? That would probably never be known.

Mike Baker reported from Seattle and Sheri Fink from New York.

The Coronavirus Outbreak >

Frequently Asked Questions and Advice

Updated April 11, 2020

When will this end?

This is a difficult question, because a lot depends on how well the virus is contained. A better question might be: “How will we know when to reopen the country?” In an American Enterprise Institute report, Scott Gottlieb, Caitlin Rivers, Mark B. McClellan, Lauren Silvis and Crystal Watson staked out four goal posts for recovery: Hospitals in the state must be able to safely treat all patients requiring hospitalization, without resorting to crisis standards of care; the state needs to be able to at least test everyone who has symptoms; the state is able to conduct monitoring of confirmed cases and contacts; and there must be a sustained reduction in cases for at least 14 days.

READ MORE 