

# Open Science Prize goes to software tool for tracking viral outbreaks

February 28 2017

---

After three rounds of competition—one of which involved a public vote—a software tool developed by researchers at Fred Hutchinson Cancer Research Center and the University of Basel to track Zika, Ebola and other viral disease outbreaks in real time has won the first-ever international Open Science Prize.

Fred Hutch evolutionary biologist Dr. Trevor Bedford and physicist and computational biologist Dr. Richard Neher of the Biozentrum Center for Molecular Life Studies in Basel, Switzerland, designed a prototype called nextstrain to analyze and track genetic mutations during the Ebola and Zika outbreaks. Using the platform Bedford and Neher built, anyone can download the source code from the public-access code-sharing site GitHub, run genetic sequencing data for the outbreak they are following through the pipeline and build a web page showing a phylogenetic tree, or genetic history of the outbreak, in a few minutes, Bedford said.

He and Neher envision the tool as adaptable for any virus—a goal to which they will apply the \$230,000 prize announced today by its three sponsors, the U.S. National Institutes of Health, the British-based charitable foundation Wellcome Trust and the U.S.-based Howard Hughes Medical Institute.

"Everyone is doing sequencing, but most people aren't able to analyze their sequences as well or as quickly as they might want to," Bedford said. "We're trying to fill in this gap so that the World Health Organization or the U.S. Centers for Disease Control and

Prevention—or whoever—can have better analysis tools to do what they do. We're hoping that will get our software in the hands of a lot of people."

For now, the tool is easy to use for Zika and Ebola. (The researchers also built a separate platform called nextflu for influenza.) But adapting the platform for other pathogens still involves a fair amount of work and technical skill, so Bedford is working with a web developer to "get that bar down so it will be easier to have this built out for other things."

By lowering the technical bar, he and Neher hope to nudge researchers to overcome another obstacle: a longstanding reluctance to share data. That is also a goal of the Open Science Prize.

## **Sharing is caring**

"Open science" supporters believe, as Bedford and Neher do, that sharing preliminary information quickly speeds discoveries, including those that could improve human health, and is therefore good for both science and society. The Open Science Prize competition aimed to stimulate the development of ground-breaking tools and platforms to make it easier for researchers and the wider public to share and find publications, datasets, code and other research outputs as well as to "generate excitement, momentum and further investment" in doing so, according to the prize sponsors.

Nextstrain "is an exemplar of open science and will have a great impact on [public health](#) by tracking viral pathogens," said Robert Kiley, who leads Wellcome's work on open research, in a statement. All of the Open Science Prize entrants "demonstrated what's possible when data and code are made open for all," he said.

Bedford and Neher were among six teams of finalists chosen in May

from 96 entries representing 450 innovators and 45 countries. In January, a public vote (3,730 votes from 76 countries, to be precise) narrowed the field to three. Bedford praised both runner-up teams as doing "really fantastic work." MyGene2 is designed to help people with rare diseases share health and genetic information with other families, clinicians and researcher worldwide. OpenTrialsFDA is aimed at making it easier to find information from clinical trials that was reported to the federal Food and Drug Administration but never published in academic journals.

For all of its cutting-edge technology, nextstrain, the winning project, belongs to a long tradition of using data visualization to understand—and intervene in—outbreaks, dating back to the 1854 London cholera outbreak. At the time, cholera, an infectious and often fatal intestinal disease, was thought to be spread by "miasma" or bad-smelling air. Dr. John Snow, the "father of modern epidemiology," the study of the causes and patterns of disease, suspected the disease was spread by contaminated water. He drew a map of public well sites and cholera cases and noted that cases clustered around a particular well.

The map, Bedford said, made an intervention—removing the handle of the Broad Street water pump—obvious.

"What we're doing with nextstrain is meant to be in this tradition," he said. "Right now it's more of a 'now-cast,' but we really want to be doing a real-time forecast of what's going on with an epidemic."

Evolutionary and computational biologists like Bedford and Neher are in the [open science](#) movement's vanguard. One reason is that their fields are the ones most concerned with outbreaks, where waiting to publish can have deadly consequences.

Real-time tracking of genetic mutations during disease outbreaks helps

scientists discern what makes viruses so severe and inform public health efforts to contain them. Being able to do so depends on researchers openly sharing the genetic sequencing data, something that not all scientists embrace in a competitive world where researchers rush to publish in prestigious journals and stake claims to discoveries.

## **Lessons from Ebola**

The seed for nextstrain sprouted while Bedford was doing postdoctoral research at the University of Michigan. He had published a paper on flu migration using data up to 2010. He found himself thinking what a pity it was that the analysis couldn't be updated as new data came out. But the fact that a paper had already been published was a disincentive for anyone to write a new paper with just a small update to the data.

From that frustration, nextflu was born. And nextflu led to nextstrain.

The devastating 2013-2016 Ebola epidemic in West Africa lent the project new urgency. Relatively early in the outbreak, researchers sequenced Ebola genomes from patients and immediately uploaded them to the public database GenBank, leading to a surge of collaboration from experts in diverse fields. The collection of shared, publically available data helped answer critically important questions as the epidemic was unfolding. It added to the confirmation that that the outbreak was being sustained by human-to-human contact, not contact with bats or other animal carriers, suggested probable transmission routes and revealed where and how fast mutations in the virus were occurring—all information crucial to both public health and medical interventions.

Even when data is shared, speed is everything in responding to outbreaks, so any tool that speeds data analysis contributes to the effort.

But despite the precedent set by the response to the Ebola epidemic,

fewer researchers have shared Zika virus genome sequences from the more recent crisis in Brazil, Central America and the Caribbean, the researchers said.

"I'm not seeing the same thing with Zika," said Dr. Gytis Dudas, a postdoctoral fellow in Bedford's laboratory who worked on many of the Ebola analyses. In part, Dudas said, the Zika virus is more difficult to sequence than Ebola, making researchers more likely to guard their rare sequences for publications.

And that, Bedford said, is "a tragedy," even as he understands that academic careers depend on publishing.

"The idea is that this nextstrain platform would provide some neutral ground with which to share data," said Bedford. "We're not trying to make a flashy paper. We just want [the data] to be on the website so people can look at the latest thing and do analyses that aren't stymied by publication practices. This kind of simple sequence sharing during outbreaks is something that if you could just push the [scientific] community a little bit, you could have some real-world impact in helping respond to epidemics."

Provided by Fred Hutchinson Cancer Research Center

Citation: Open Science Prize goes to software tool for tracking viral outbreaks (2017, February 28) retrieved 7 July 2024 from <https://phys.org/news/2017-02-science-prize-software-tool-tracking.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.