# 🦫 Healthcare industry BW

"Open science" – a successful recipe for combating viruses

Danger identified, danger averted – the better the dynamic changes of the viral genome are understood, the better the spread of the pathogens can be predicted, thus enabling more efficient countermeasures to be taken. A team of researchers from Tübingen and Seattle is developing an open source online platform that maps the evolution of viruses and identifies dangerous developments.
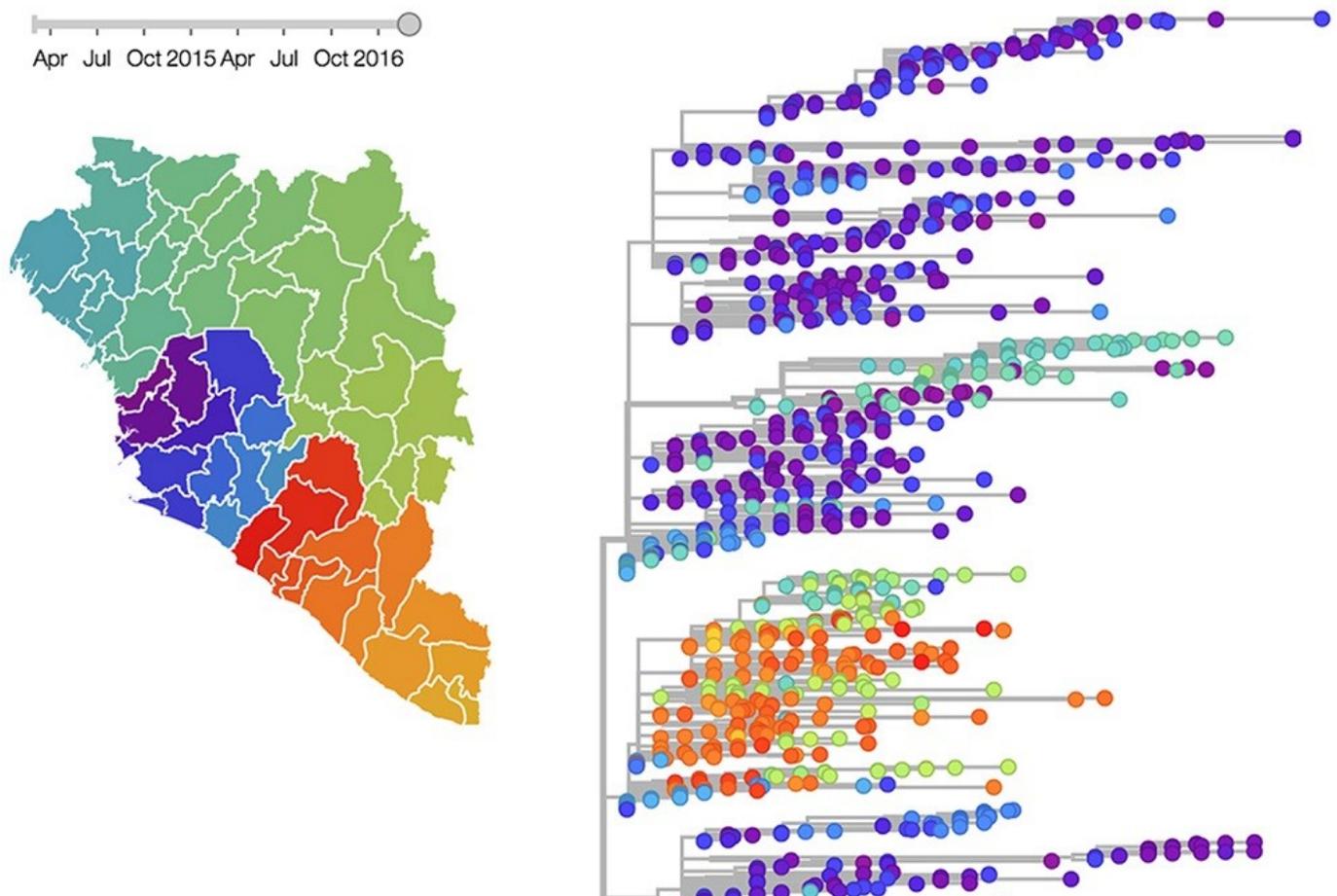
Phylogenetic tree of Ebola viruses and the geographic distribution of Ebola infections.
© R. Neher / T. Bedford

When a virus epidemic breaks out, it is always a race against time. The sooner infected people are identified, treated and, if necessary, isolated, the sooner the spread of the virus can be curbed. However, health authorities can only take appropriate action when comprehensive, current

epidemiological information about the virus is available. Researchers from the Max Planck Institute for Developmental Biology in Tübingen have developed a useful tool for providing this type of information.

Dr. Richard Neher and his team are working with scientists from Dr. Trevor Bedford's team at the Fred Hutchinson Cancer Research Center in Seattle on the development of an open source online platform to collect, analyse and present comprehensive information about current virus strains, their geographic distribution and transmission routes. They are developing special new algorithms and visualisation systems.

Viruses multiply extremely quickly and have a high mutation rate. And this is what the researchers from Tübingen and Seattle are using to their advantage. The software they have developed has the capability to determine the relationship between viruses based on spontaneous mutations, which occur relatively frequently. The software compares genome sequences from a large number of viruses and uses this information to construct phylogenetic trees that can be used to deduce the viruses' evolutionary development and future spread. New information is constantly being added.

"Every tree looks different. Viruses like HIV which have a far more complex transmission route than influenza viruses, also have a completely different dynamic. In contrast to HIV, influenza viruses can easily be transmitted by inhalation of airborne droplets. Such differences are reflected in the pattern of the phylogenetic trees," says Neher. Despite the different patterns, the trees produced by the new software all enable the user to easily see which mutations spread and how fast they spread. "Viruses do not evolve towards something, but away from something. This means that they mutate to evade immunisations. Resistances to drugs develop," says Neher. Presenting these relationships helps researchers develop vaccines. With flu vaccines, which have to be newly designed every year, the software will make it a lot easier to identify the influenza virus strains that are most likely to spread and cause illness each flu season, and subsequently to adapt the vaccine to these strains.

Creating structures to collect and share data quickly

The online system is designed such that all data, regardless of where in the world they are entered, are immediately incorporated into the analysis, and hence into the phylogenetic tree, so that they are available in real time to researchers around the world. "The key issue is being able to access the data as quickly as possible. The biggest challenge is the logistics," says Neher. However, the results so far have been quite positive. "In the case of the Ebola virus, the RNA sequences are available around a week after sampling. As far as the influenza season is concerned, we usually get new sequence information about the viruses that cause disease during this particular period at fortnightly intervals." Even in the case of the current Zika epidemic in Brazil, the team receives data relatively early thanks to mobile laboratories across the country. The analysis only takes around an hour and can then be made available online.

However, the information provided by the software is only of benefit if the key players follow up on it, especially health centres in the country experiencing the virus epidemic. In order to raise awareness among those dealing with the epidemic and facilitate the use of the online portal, Neher and his colleagues have run workshops at national flu centres. This is also of benefit for the scientists themselves, because they can find out during the workshops what the practical difficulties are and where they need to improve their online platform.

"Open science" as a success

The researchers hope that open access and exchange of data will help improve the dynamics of the

online platform and ensure its success. At the same time, Neher is well aware of the problems associated with open science platforms. "Sometimes, researchers choose not to disclose information because they want to publish it. We are therefore working on finding a solution so that scientists can provide data for use on our platform and still be able to publish them," says Neher. He is extremely keen to publish data and make data and results available to the wider public in order to improve epidemiological predictions. Neher also believes that this would have a macroeconomic benefit based on this rough estimate: "If our predictions could be used to improve the efficacy of influenza vaccines by two percent or so, we would have around one million fewer cases of illness and 10 million fewer days of sickness."

Even large American institutions see the value of open data exchange. The American Wellcome Trust, the National Institutes of Health (NIH) and the Howard Hughes Medical Institute have launched the "Open Science Prize" aimed at "unleashing the power of open content and data to advance biomedical research and its application for health benefits". The team led by Neher and Bedford were among the six winners of the first phase of the Open Science Prize in 2016. In May 2016, they were awarded US$ 80,000 which they will use for optimising the online platform before submitting it to the second phase of the prize. In parallel, Neher is thinking of expanding the database to other pathogens, including animal viruses and bacteria. "In bacteria, there are many things that are not linked to mutations. Bacteria exchange entire plasmids, which presents us with entirely new technical challenges," says Neher.
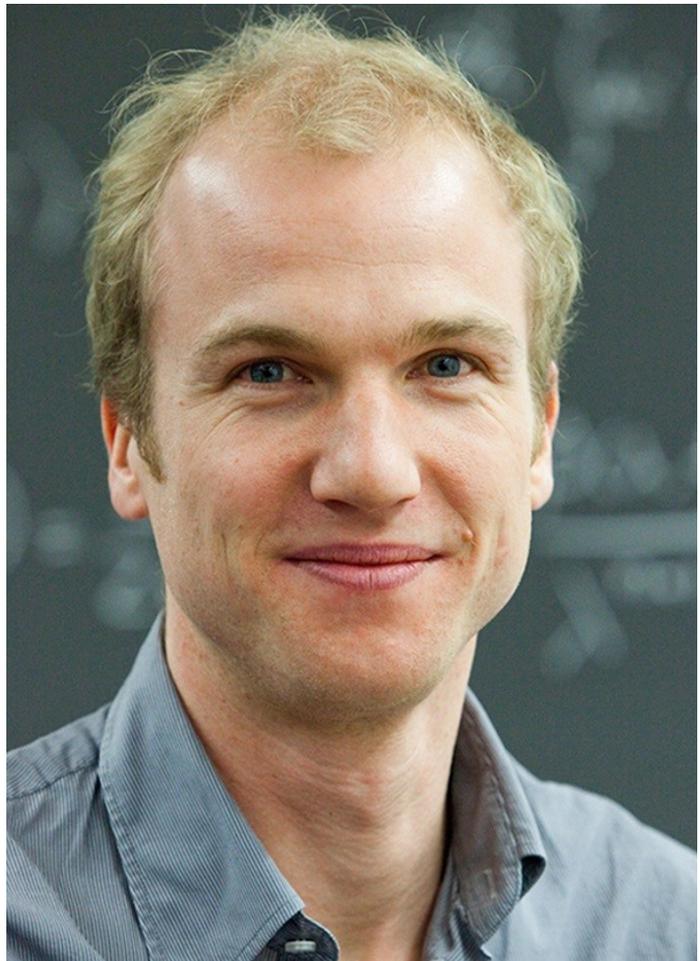
Dr. Richard Neher is using his ERC Starting Grant to develop the open science platform, amongst other things.
© Jörg Abendroth/Max Planck Institute for Developmental Biology

Article

Further information

Max Planck Institute for Developmental Biology
Dr. Richard Neher
Spemannstr. 35
72076 Tübingen
Tel.: +49 (0)7071 601-1345
E-mail: richard.neher(at)tuebingen.mpg.de

The article is part of the following dossiers

Data mining: new opportunities for medicine and public health

Human infectious diseases: new threats

bioanalytics · infectious diseases · prevention · sequencing · virus · evolution · software · database · vaccine · basic research · influenza · DNA analysis · data mining