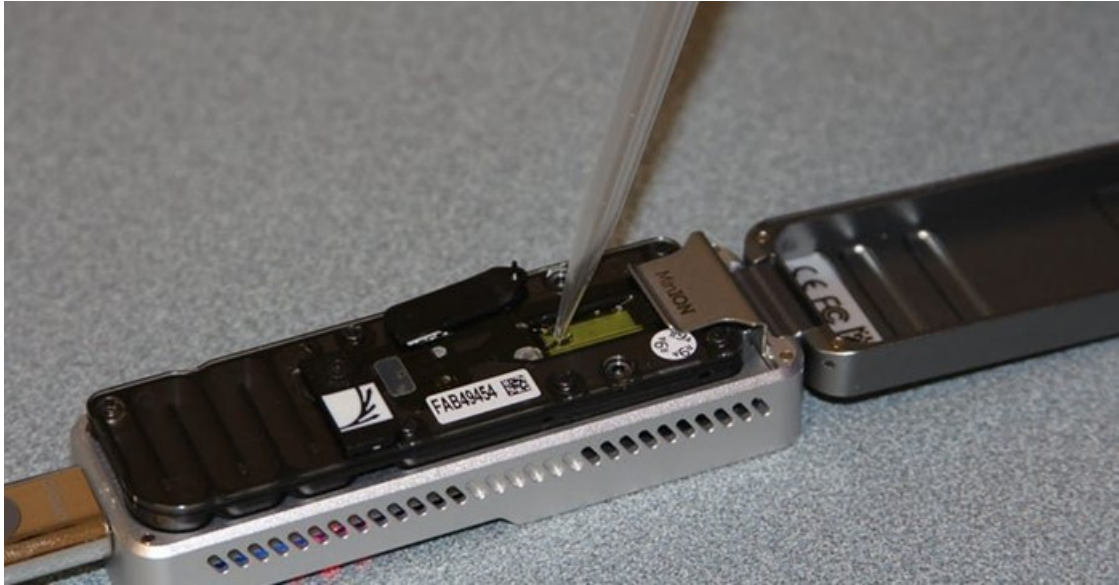


DNA sequencing can help fight epidemics - but there are privacy risks

By [Liam Shaw & Nicola C Sugden](#)

14 Jan 2019

The Democratic Republic of Congo is battling an [Ebola outbreak](#). As is the case with any disease caused by pathogenic viruses - like Zika or influenza - Ebola spreads dangerously and unpredictably. This makes tracking the movement of viruses around the world a major challenge.



A portable DNA sequencer in action. UGA CAES/Extension/Flickr

Researchers have increasingly turned to DNA sequencing to help identify and track these sorts of diseases. They use portable DNA sequencers, which are the size of a USB and can be easily carried for use in the field. One such sequencer, the [MinION](#) from Oxford Nanopore Technologies [was used](#) during the 2016 Zika virus epidemic in Brazil. It's also being used [to track](#) the DRC's Ebola outbreak.

Some researchers hope it will soon be possible to combine sequencing data collected in this way with other information to tell us even more about disease outbreaks. Integrating different kinds of data into a global infectious disease surveillance system that continuously scans for new epidemics might make it possible to detect outbreaks and sequence viruses as they emerge, allowing public health responses to be suggested in real time.

There's no doubt these efforts are driven by good intentions. But, as we argue in [our new research](#), this technology – which supporters hope will become increasingly available to [members of the public](#) – could have serious privacy implications.

Metagenomic data – the kind that could be collected on a sequencer such as MinION or others such as the Chan Zuckerberg Initiative’s new platform IDSeq – contains an enormous amount of information about who we are and how we live. In combination with other widely available information someone could potentially use that data to work out where you live, or with whom you have a close relationship.

The reality is that, as improvements in data analysis methods allow us to extract new insights from old data (or [de-anonymise anonymised data](#)), it’s impossible to be absolutely sure what the potential uses of data will be.

Signing away your data

Imagine having an app on your smart phone that allows you to analyse samples from the world around you. You could use it to sequence your pet cat’s DNA, or to figure out whether the mould growing in your shower is dangerous.

Sound far-fetched? It’s not. The technology required is already here. For example, the Chan Zuckerberg Initiative recently announced [IDSeq](#), a new platform and database for infectious disease surveillance where registered users [can upload](#) their metagenomic sequencing data to have it analysed for free.

There’s just one catch, as there would be with any sequencing app: you have to sign over permissions to the data. Most people will do this unthinkingly. Author Jamie Susskind [has called](#) this pervasive and common arrangement “the data deal”: people accept whatever a company asks so they can use an app or product, and worry about the implications later.

This is the case with IDSeq. Initially enthusiastic researchers became [concerned](#) when they realised the platform’s terms and conditions contained a clause granting the Chan Zuckerberg Initiative “perpetual” permission to “use”, reproduce, distribute, display and create derivative works” from the data.

The current justification given for this clause is that it’s intended to permit users’ research data to be used for improving IDSeq. However, in principle the data could later be shared with [“any third party that purchases”](#) part of the assets or organisation.

A world of information

It could contain not only the DNA you wanted to sequence, but also DNA from your fingers when you loaded the sample, from the bacteria on your skin from the last person you hugged, or from the gardens your cat visited last night. In short, that data contains vital information about your [microbiome](#) – the vast collection of microorganisms that live on and in our bodies. And your microbiome can tell someone an awful lot about you.

As we learn more about our microbiomes, we are beginning to understand how much they are personalised. Even if we could filter out the human DNA sequences from datasets, our microbiomes could [theoretically](#) still be used to identify us.

The microbiome contains information not only about our lifestyles, like our diet and drug intake, but also our social relationships, such as who we live with. That’s a lot of information to work with, in a world where we already share a great deal of data about ourselves via platforms like Facebook and Instagram, or personal fitness trackers. This data could feasibly be merged with metagenomic data, making it even more powerful.

There are ever more surprising examples of incidental data being used in dramatic and unexpected ways that are far removed from the original reasons for collecting it. Data from a murder victim’s Fitbit was [used](#) to convict her killer. And data from users of the fitness app Strava inadvertently [revealed](#) the location of secret US army bases.

There is every reason to believe that data from portable sequencers collected primarily for disease surveillance would

contain information that could be used in similarly surprising, and concerning, ways. Metagenomic sequencing data is highly personalised. It contains implicit information about who we interact with and where we go, which makes it commercially valuable.

These concerns shouldn't (and won't) stop portable sequencers being used for infectious disease surveillance. Corporations and governments will promise great benefits from the use of this technology. For example, the IDSeq privacy notice justifies data collection by [appealing](#) to "legitimate interest in investigating and stopping the spread of infectious diseases and promoting global health".

We need to continue scrutinising these organisations to make sure we understand exactly what's being done with our data. The consequences of widespread portable sequencing, like emerging infectious diseases themselves, will be highly unpredictable.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

ABOUT THE AUTHOR

Liam Shaw, *University of Oxford* and Nicola C Sugden, *University of Manchester*

For more, visit: <https://www.bizcommunity.com>