Scientists pioneer new way of finding cancer-causing germs

October 22 2019

Credit: CC0 Public Domain

Scientists at the University of East Anglia are pioneering a new way of finding the bacteria and viruses associated with cancer.

Some infections have already been linked with cancers including stomach cancer and cervical cancer.

New research published today in *Genome Biology* shows how genomic data collected from tumours could be sequenced to reveal other associated bacteria and viruses.

It is hoped that the work could help detect cancer-causing germs, and one-day even lead to more cancer vaccines like the HPV vaccine.

Lead researcher Dr. Dan Brewer, from UEA's Norwich Medical School and a Visiting Worker at the Earlham Institute (EI), said: "We already know of some strong associations between infections and cancer. For example the presence of Helicobacter pylori in the digestive tract can lead to stomach ulcers and is associated with stomach cancer.

"Meanwhile some types of the HPV virus can cause cervical cancer.

"Human tissue is increasingly being whole genome sequenced as we transition into an era of genomic medicine.

"When tumour samples are whole genome sequenced, DNA from any pathogens present will also be sequenced, making it possible to detect and quantify pathogens.

"This gives us a fantastic opportunity to collect data that will help us find new associations between bacteria and viruses and different types of cancer."

The research team produced computer simulations to test multiple approaches before running tests on real cancer sequence data where an array of interesting pathogens were detected.

Postgraduate researcher Abraham Gihawi, also from UEA's Norwich Medical School, said: "There are already a wealth of scientific tools to classify organisms from their genome sequence data. This new research looks at each of the key computational steps involved in conducting this on human tissue sequencing data.

"We collected computer scripts that stitch together the highest performing tools to apply to large collections of sequencing data.

"We used computer simulations and then tested our approaches on some real cancer sequence data.

"For the simulations, we produced 'mock' genomes containing mostly human sequences and spiked them with very small numbers of pathogenic sequences which is what we expect to see in cancer sequences.

"Since we knew exactly what pathogens, and how much of them, are in each simulated genome, we could test how well each approach performs.

"We then tested our method out on some real cancer sequencing data where we know what we expect to see.

"As well as finding pathogens that are known to be associated with cancer, we also detected many more. We are only just beginning to scratch the surface on the role that these other pathogens may play in the development of cancer."

Dr. Brewer added: "There is already an HPV vaccine which is thought to prevent around 70 per cent of cervical cancer cases. We hope that by identifying bacteria and viruses associated with other cancers, new vaccines could be developed in the future."

More information: Abraham Gihawi et al. SEPATH: benchmarking the search for pathogens in human tissue whole genome sequence data leads to template pipelines, *Genome Biology* (2019). DOI: 10.1186/s13059-019-1819-8

Provided by University of East Anglia