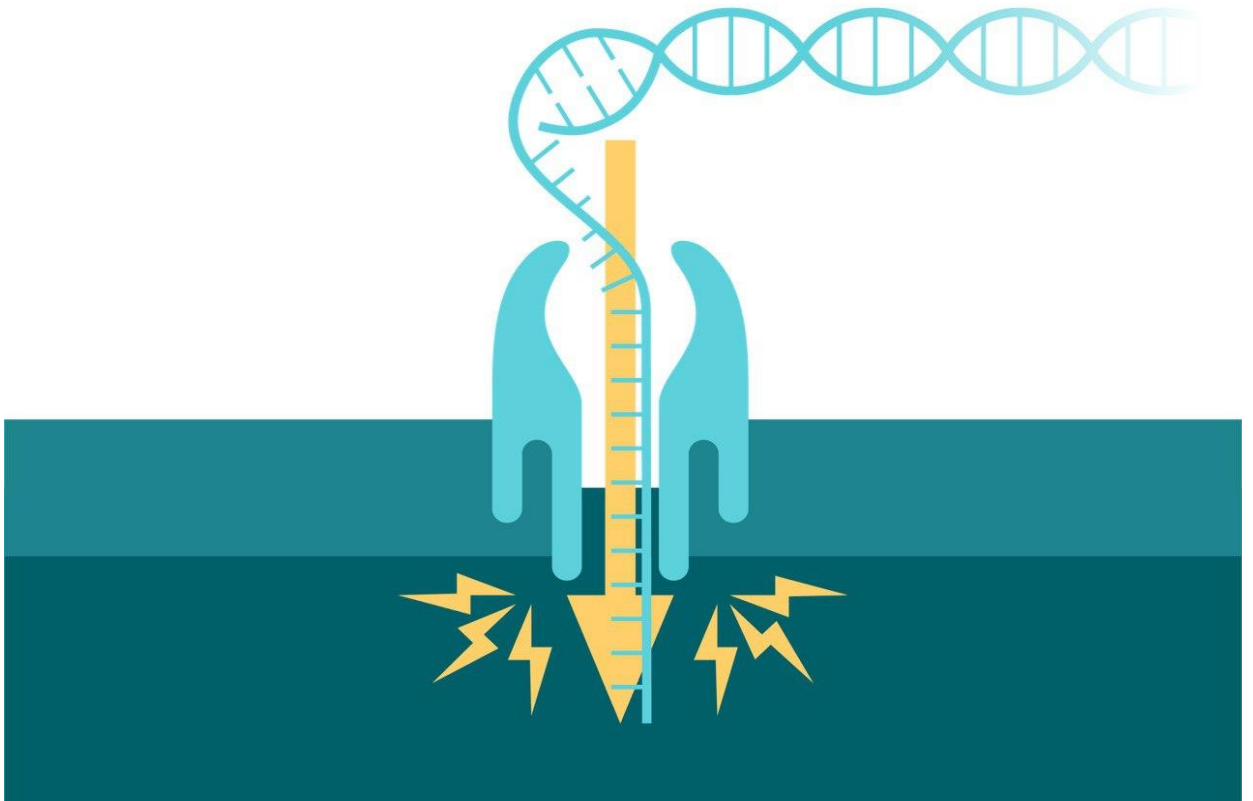# EXTRACTING BIOLOGY FROM NANOPORE SIGNALS WITH ARTIFICIAL INTELLIGENCE

Ion current signals are emitted from a nucleic acid molecule passing through a nanopore channel.

Sequencing of DNA and RNA is one of the central technologies in genomic research. The Oxford Nanopore Technologies™ sequencing method can profile molecules that are much longer compared to other sequencing technologies, which can resolve genomic and transcriptomic regions. Nanopore sequencing emits ion current signals during the sequencing process that can provide further insights on epigenomic or epitranscriptomic changes in a single assay. Due to the highly complex nature of the signals, artificial intelligence methods that use machine-learning are employed to extract biological information from these signals. In recent years, the development of machine-learning-based tools for nanopore ion current signals analysis has improved the accuracy of Nanopore sequencing and expanded the types of biological profiling, providing a simple approach for a fuller picture of the biology behind both DNAs and RNAs.

Based on their experience in analysing nanopore sequencing data at the Genome Institute of Singapore (GIS), Ms. Yuk Kei Wan, Mr. Christopher Hendra, Dr. Ploy Pratawanich, and Dr. Jonathan Göke summarised the development of machine-learning-based nanopore ion current signal analysis software in a review published in *Trends in Genetics*. The review explores and categorises the machine-learning approaches used in the tools for associating ion current signals to nucleotide sequences and extracting DNA and RNA features from ion current signals.

In addition to the review, the team also provided a curated list of nanopore analysis software on the awesome-nanopore GitHub respository.

> "Artificial intelligence research has developed many approaches to analyse data that is similar to the output from Nanopore sequencing. Over the last years we have seen a growth of new methods often adopting and extending these approaches to extract new biological insights from Nanopore signal data. Some of these are actually contributed by scientists from the Genome Institute of Singapore, so we are very excited to provide this timely overview of machine learning applications in this fast growing area of genomics."
>
> *Dr Jonathan Göke, Group Leader of the Laboratory of Computational Transcriptomics at GIS*