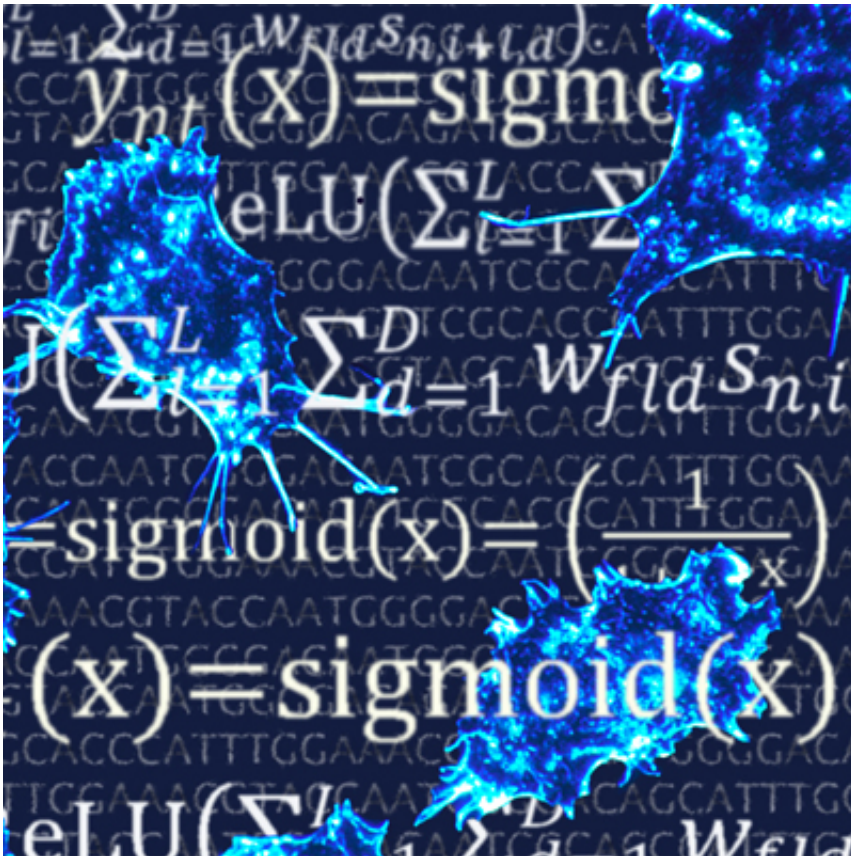


New deep learning technique offers a more accurate approach to single-cell genomics

April 12 2017



A new 'deep learning' method, DeepCpG, has been designed by researchers at the Wellcome Trust Sanger Institute, the European Bioinformatics Institute and the Babraham Institute to help scientists better understand the epigenome – the biochemical activity around the

genome. Reported today in *Genome Biology*, DeepCpG leverages 'deep neural networks', a multi-layered machine learning model inspired by the brain, and provides a valuable tool for research into health and disease.

As a result of projects like 1000 Genomes, scientists now have a 'book' of the human genome divided up into chapters and annotated in parts. However, to fully understand how life works, scientists need to decipher both the genome – the set of instructions repeated in every cell – and the epigenome, the part that varies wildly between cells.

To better understand how DNA sequences relate to biological changes, the genomics community is turning to artificial [neural networks](#) – a class of machine learning methods first introduced in the 1980s and inspired by the wiring of the brain. More recently, these models have been rebranded as 'deep neural networks', which form the field of deep learning.

Scientists have leveraged the capacity of deep learning to fill in the gaps in single-cell genomics, an emerging technology that offers a close-up view on epigenetics.

A new technique, DeepCpG, has been designed to help scientists learn about the connections between DNA sequences and DNA methylation – a biochemical modification of the genome sequence that can act like an off-switch for individual genes. Methylation plays a key part in important biological processes, including cell development, ageing and cancer progression.

The new method uses genomic and epigenomic data to make predictions about DNA methylation in [single cells](#). This is important because current technologies provide incomplete information about this. With DeepCpG, researchers can obtain a more complete picture of DNA methylation. The model can also be used to obtain new biological insights, for

example on the connection between the DNA sequence and methylation.

"DeepCpG actually learns meaningful features in a data-driven manner. It has major advantages over previous methods, including the ability to more accurately predict DNA methylation and to study intercellular differences. By studying the wiring of the learnt network, we can understand how the biology of DNA methylation works. This has allowed us to recover known DNA sequence motifs that are important for methylation changes, as well as to discover new motifs, which are the starting point for future studies," says Christof Angermueller, PhD candidate at EMBL-EBI.

"We have demonstrated that DeepCpG enables us to accurately predict and analyse DNA methylation in single cells. However, DeepCpG is just one example of how we can apply deep learning to genomics and single-cell technologies. It is exciting to see the versatile applications deep learning has already found in genomics. I am looking forward to seeing more deep learning techniques come online. I believe it will make a big difference to how we study biology and has the potential to yield new answers about how life works," says Dr Oliver Stegle, Group Leader at EMBL-EBI.

"Single cell epigenomics methods provide exciting insights into cell heterogeneity in development, ageing and disease; however if you are just dealing with two genomes in a single cell, bits of information are often lost during the experiment. This new method recognises patterns of the epigenome in single [cells](#) and then reconstructs lost information, returning a data-rich single cell epigenome," says Professor Wolf Reik from The Babraham Institute and Associate Faculty member at the Wellcome Trust Sanger Institute.

"Deep learning is now the state-of-the art in many fields. We are

exploring its utility for making sense of large scale biological data. Pioneering studies, such as the one by Angermueller and colleagues, prove that there is lot to be gained by using [deep learning](#) methods in computational biology," says Dr Leopold Parts, Group Leader at the Sanger Institute.

More information: Christof Angermueller et al. DeepCpG: accurate prediction of single-cell DNA methylation states using deep learning, *Genome Biology* (2017). [DOI: 10.1186/s13059-017-1189-z](https://doi.org/10.1186/s13059-017-1189-z)

Provided by Wellcome Trust Sanger Institute

Citation: New deep learning technique offers a more accurate approach to single-cell genomics (2017, April 12) retrieved 18 April 2024 from <https://medicalxpress.com/news/2017-04-deep-technique-accurate-approach-single-cell.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.