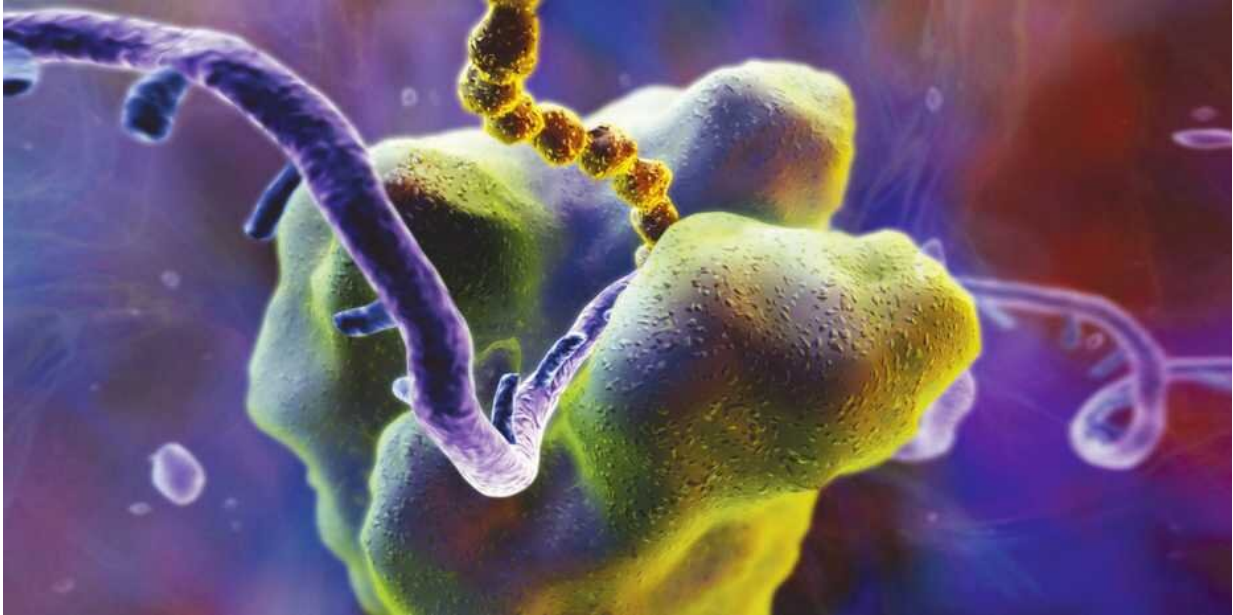


Binding sites for protein-making machinery

August 27 2020, by Fabio Bergamin



Ribosomes are the cell's protein-making machinery. They read the genetic information encoded in messenger RNA (violet) and translate it into proteins (yellow). Credit: Science Photo Library

ETH Zurich researchers can predict how tightly a cell's protein synthesis machinery will bind to RNA sequences—even when dealing with many billions of different RNA sequences. This binding plays a key role in determining how much of a specific protein is produced. The scientists are developing their prediction model using a combination of synthetic biology experiments and machine learning algorithms.

Genome sequencing of bacteria, plants and humans has become a routine process, yet the genome still poses many unanswered questions. One of these concerns the sites on messenger RNAs (mRNAs) that ribosomes—the cellular structures responsible for protein synthesis—bind to in order to translate genetic information. Currently, the function of these ribosome binding sites is only partly understood.

An interdisciplinary team of researchers from the Department of Biosystems Science and Engineering (D-BSSE) at ETH Zurich in Basel has now developed a new approach that, for the first time, makes it possible to obtain detailed information on an incredibly large number of these binding sites in bacteria. The new approach combines experimental methods of synthetic biology with machine learning.

Precise control over protein production

Ribosome binding sites are short RNA sequences upstream of a gene's coding sequence. In the past, biotechnologists also developed synthetic binding sites. The ribosomes bind extremely well to some of these, and less well to others. The tighter ribosomes are able to bind to a specific variant, the more often they translate the respective gene and the greater the amount of the corresponding protein they produce.

Biotechnologists who use bacteria to produce chemicals of interest such as pharmaceuticals can influence the amount of involved proteins in the cell through their choice of ribosome binding sites. "Exerting this kind of control is particularly important and helpful when incorporating complex gene networks comprising multiple proteins at the same time. The key here is to establish an optimal balance amongst the different proteins," says Markus Jeschek, senior scientist and group leader at D-BSSE.

An experiment with 300,000 sequences

Together with ETH professors Yaakov Benenson and Karsten Borgwardt and members of the respective groups, Jeschek has now developed a method to determine how tightly ribosomes bind to hundreds of thousands or more RNA sequences in a single experiment. Previously this was only possible for a few hundred sequences.

The ETH researchers' approach harnesses deep sequencing, the latest technology used to sequence DNA and RNA. In the laboratory, the scientists produced over 300,000 different synthetic ribosome binding sites and fused each of these with a gene for an enzyme that modifies a piece of target DNA. They introduced the resulting gene constructs into bacteria in order to see how tightly the ribosomes bind to RNA in each individual case. The better the function of the binding site, the more enzyme is produced in the cell and the more rapidly the target DNA will be changed. At the end of the experiment, the researchers can read this change together with the binding [site](#)'s RNA sequence using deep sequencing.

Universally applicable approach

Since 300,000 represents only a small fraction of the many billions of theoretically possible [ribosome](#) binding sites, the scientists analyzed their data using machine learning algorithms. "These algorithms can detect complex patterns in large datasets. With their help, we can predict how tightly ribosomes will bind to a specific RNA sequence," says Karsten Borgwardt, Professor of Data Mining. The ETH researchers have made this prediction model freely available as software so that other scientists can make use of it, and they will soon be introducing an easy-to-use online service as well.

The approach developed by the scientists is universally applicable, Benenson and Jeschek emphasize, and the team is planning to extend it to other organisms including human cells. "We're also keen to find out how [genetic information](#) influences the amount of [protein](#) that is produced in a human cell," Benenson says. "This could be particularly useful for genetic diseases."

More information: Simon Höllerer et al. Large-scale DNA-based phenotypic recording and deep learning enable highly accurate sequence-function mapping, *Nature Communications* (2020). [DOI: 10.1038/s41467-020-17222-4](#)

Provided by ETH Zurich

Citation: Binding sites for protein-making machinery (2020, August 27) retrieved 27 April 2024 from <https://phys.org/news/2020-08-sites-protein-making-machinery.html>

<p>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.</p>
--