

Prediction of susceptible genes associated with diabetes risk

April 20 2016

The Institute for Infocomm Research (I²R) is a member of the Agency for Science, Technology and Research (A*STAR) family and is Singapore's largest ICT research institute. Established in 2002, our vision is to power a vibrant and strong infocomm ecosystem in Singapore. We seek to foster world-class infocomm and media research and develop a deep talent pool of infocomm professionals to power a vibrant knowledge-based Singapore.

Give me a drop of blood and I can tell you your genetic risk of developing diabetes within the next five to ten years! No, this is not a science-fiction project. It is a scientific project that could potentially revolutionize the field of diagnostic medicine by leveraging the strengths of two techniques—Genome Wide Association Studies (GWAS) and Machine Learning—from two very different disciplines. This work is still at the research level and it will be at least a few years before it can be implemented in clinics.

After the Human Genome Project finished in 2002, GWAS gained significant momentum and became much more affordable. Although genetics is difficult for a layman to understand, it can be thought of as analogous to the English language. The English alphabet has 26 letters, whereas genes have only four letters or 'bases'—A(adenine), T(thymine), G(guanine) & C(cytosine). Our genome is entirely made up of these four letters, albeit in different lengths and combinations. To extend the analogy, similar to the Autocorrect (in MS Word) detecting spelling mistakes in your text, GWAS tries to pick up on spelling mistakes in our

genome, called Single Nucleotide Polymorphisms (SNPs). GWAS analyzes the effects, both good and bad, that these SNPs have on a person. While some SNPs might be benign and result in curly hair or sharper nose, others might have a detrimental effect and cause cancer or other diseases. To put the scale of data to be analyzed in perspective—a typical genome has anywhere between 600,000 to 1 million SNPs! Adding to the complexity, most of the phenotypic changes are caused not just by a single SNP but interaction of many SNPs.

While traditional statistical methods aren't adept at handling interaction effects among the variables involved, machine learning techniques can process these interactions with ease. By employing machine learning techniques, the SNP(s) that are associated with a susceptibility to diabetes are identified and the genetic risk of an individual developing diabetes can be predicted. Note however, that this estimate is based solely on genetic information and does not take lifestyle factors into account.

Genes do play a role in diabetes, but lifestyle choices are also important. The following two scenarios show that genes aren't the only contributing factor in developing diabetes:

1. A person can have a genetic mutation that may make them susceptible to diabetes, but by following a healthy lifestyle (exercise/diet) they may not develop diabetes.
2. While identical twins have the same genome, according to the American Diabetes Association, when one twin develops type 1 or type 2 diabetes, the other twin has at most a 50 or 75 per cent chance of developing type 1 or type 2 diabetes, respectively.

This clearly suggests that although a person's susceptibility to the disease is determined by genetics, [it is something in their environment that actually triggers it.](#)

This work is being validated on a real patient dataset obtained from the School of Public Health, the National University of Singapore (NUS).

There are plenty of follow-up tasks we can perform to cater to different applications. Firstly, now that the pipeline for this [diabetes](#) GWAS study is built it can easily be extended to predict other diseases like stroke, cancer(s), etc. which could save many lives. Secondly, from a therapeutics perspective, drugs can be designed to specifically target the genes responsible for malign SNPs.

So let's look forward to a future where we can predict the [genetic risk](#) of a disease five to ten years in advance, provide target medication and enhance the lifespan of mankind.

Provided by Agency for Science, Technology and Research (A*STAR), Singapore

Citation: Prediction of susceptible genes associated with diabetes risk (2016, April 20) retrieved 3 July 2024 from <https://medicalxpress.com/news/2016-04-susceptible-genes-diabetes.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>
--