Medical Xpress

Study reveals broad 'genetic architectures' of traits and diseases

August 13 2018

A depiction of the double helical structure of DNA. Its four coding units (A, T, C, G) are color-coded in pink, orange, purple and yellow. Credit: NHGRI

Scientists at Johns Hopkins Bloomberg School of Public Health have developed a powerful method for characterizing the broad patterns of genetic contributions to traits and diseases. The new method provides a "big picture" of genetic influences that should be particularly helpful in designing future genetic studies and understanding potential for genetic risk prediction.

The scientists, in a study published on Aug. 13 in the journal *Nature Genetics*, mined existing data from genetic studies and used novel statistical techniques to obtain estimates of the numbers of DNA variations that contribute to different physical traits and diseases, including height, BMI, childhood IQ, Alzheimer's disease, diabetes, heart disease and bipolar disorder.

"In terms of practical results, we can now use this method to estimate, for any trait or disease, the number of individuals we need to sample in future studies to identify the majority of the important genetic contributions," says study senior author Nilanjan Chatterjee, Ph.D., the Bloomberg Distinguished Professor in the Department of Biostatistics.

Affordable DNA-sequencing technology became available around the turn of the millennium. With it, researchers have performed hundreds of genome-wide association studies (GWAS) to discover DNA variations that are linked to different diseases or traits. These variations—changes in DNA "letters" at various sites on the genome—are called single nucleotide polymorphisms (SNPs). Knowing which SNPs are linked to a disease or trait can be useful in gaining biological understanding about how diseases and other traits originate and further progress.

There is also enormous interest in using genetic markers to develop risk-scores that could identify individuals at high or low risk for diseases and then use the information to develop a "precision medicine" approach to disease prevention through targeted interventions.

"Depending on their sample sizes, previous genome-wide association studies have uncovered a few SNPs or many for any given disease or trait," Chatterjee says. "But what they generally haven't done is reveal the overall genetic architectures of diseases or traits—in other words, the likely number of SNPs that contribute and the distributions of their effect sizes."

Chatterjee and his colleagues developed statistical tools to infer this overall architecture from publicly available GWAS data. They then applied these tools to 32 GWAS datasets covering 19 quantitative traits and 13 diseases.

The findings show that what is known about many traits represents the "tip of the iceberg." An individual trait could be associated with thousands to tens of thousands of SNPs, each of which has small effect, but which cumulatively make a substantial contribution to the trait variation. Intriguingly, they found that traits related to mental health and ability, such as IQ, depression and schizophrenia, appear to be the most "polygenic" in that they are influenced by the largest number, on the order of tens of thousands, of SNPs, each with tiny effects.

"For the traits we analyzed related to mental health and cognitive ability, there is really a continuum of effect sizes, suggesting a distinct type of genetic architecture," says Chatterjee, who has a joint appointment in Johns Hopkins Medicine's Department of Oncology.

By contrast, the analysis suggested that common chronic diseases such as heart disease and type-2 diabetes typically are influenced by fewer, although still a large number—on the order of thousands—of SNPs, most of which have small effects although a sizeable group "stick out" for their stronger effects.

Knowing the approximate genetic architecture of a disease or trait allows

scientists to predict how informative any new GWAS for that trait or disease will be, given the sample size. For example, projections in the study suggest that for most traits and diseases, such as heart disease and diabetes, the point of diminishing return for GWAS only starts after a sample size reaches several hundred thousand. For psychiatric diseases and cognitive traits, with their "long-tail" distributions of gene effects, diminishing returns usually won't kick in until sample sizes are even larger, i.e., in the millions, Chatterjee says. These results have implications for how useful genetic risk prediction models could be for different diseases depending on the sample size achievable for future studies.

"Our approach at least provides the best available "road map" of what is needed in future studies," Chatterjee says.

Some complex diseases are too rare to lend themselves to sample sizes anywhere near the required number in the foreseeable future, he notes. Their genetic underpinnings may therefore remain murky, with a large-scale national and international consortium effort needed to build larger GWAS.

"Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits" was written by Yan Zhang, Guanghao Qi, Ju-Hyun Park and Nilanjan Chatterjee.

More information: Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits, *Nature Genetics* (2018). DOI: 10.1038/s41588-018-0193-x , www.nature.com/articles/s41588-018-0193-x

Citation: Study reveals broad 'genetic architectures' of traits and diseases (2018, August 13) retrieved 4 May 2024 from https://medicalxpress.com/news/2018-08-reveals-broad-genetic-architectures-traits.html