HL107

The last straw did not break the camel's back: searching for causative variants of polygenic disease disregards individual predisposing genomic differences.

Yana Bromberg 1, Peter Kahn1 and Burkhard Rost 2

1 Rutgers University, USA.

2 TU Munchen, Germany

Abstract

With the advent of genomic sequencing, the number of studies designed to search for disease-causing genes or variants has grown exponentially. By far the biggest fraction of these approaches looks for common patterns in genomic variation of affected individuals, which is absent in healthy controls. Recent works have taken into account the specific variation-mediated functional disruption patterns of the affected genes. These approaches work because diseases are, arguably, extreme phenotypic variations and are often attributable to one or a few severely functionally disruptive variants. However, finding the often missing heritability requires establishing the genome-wide patterns of functional disruption, without attributing disease to a specific (small) set of variants or genes.

Here we show that a large fraction of exonic non-synonymous variants in genomes of healthy individuals are functionally non-neutral; i.e. at least a fifth of the variants in a given exome contribute to the specific differences between people. Our findings, thus, suggest a genomic basis of the different non-disease phenotypes. The fact that we remain healthy in spite of these deficiencies is evidence of our species robustness in a certain range of biochemical functionality. In individuals affected by (not obviously Mendelian) diseases, the specific functional backgrounds contribute to pathogenicity mechanisms. Distinguishing between the "combinatorial loads" of the non-neutral variants in relevant pathways of disease affected and healthy individuals will likely contribute to our understanding the genetic mechanisms of complex disease. Unfortunately, our findings indicate that, due to the limitations of our experimental (wet-lab) resolution, we are often unable to recognize the biochemical changes imposed on a specific gene/protein function by a particular mutation. Here, the use of computational techniques is key to further progress. Additionally, we show that common variants in healthy individuals, on average, affect protein function as much as the rare SNPs. Thus, both types of variation should be considered equally "suspect" in disease.

Moving forward in our understanding of disease requires novel ways of interpreting genomic patterns. This work highlights the importance of accounting for the overall functional baseline (the proverbial hay stack on the back of the camel) prior to looking for culprits of disease (the last straw). Moreover, we illustrate the value of computational techniques in determining important trends in the data, which are often invisible to standard experimental methods of investigation.