Research sheds new light on heritability of disease

January 16 2014

A group of international researchers, led by a research fellow in the Harvard Medical School-affiliated Institute for Aging Research at Hebrew SeniorLife, published a paper today in *Cell* describing a study aimed at better understanding how inherited genetic differences, or variants, predispose certain individuals to develop diseases such as type 2 diabetes. The study integrated computational methodology with experimentation to address and prove underlying genetic causes of type 2 diabetes. In principle, the new methodology can be applied to any common disease, including osteoporosis, Alzheimer's disease and cancer. The hope is that with better understanding of how DNA functions in these individuals, new treatments will follow.

Since completion of the Human Genome Project in 2003, researchers have been working to discover how genes contribute to disease. The question remains why some individuals are more at risk than others to develop certain diseases when factors such as age, gender and life-style are equal.

A small percentage of DNA contain the coded sequence that produces proteins necessary for cell growth and function. However DNA that lies outside of these coding regions play an essential role in turning genes on and off. By understanding how these regulatory regions work in concert with one another, we may identify targets for future therapies.

The method developed and tested by this study tracks patterns within regulatory regions in a number of species close or distant to humans. If a

pattern of variants in these non-coding regions is present in many species, it is likely to serve a very important function.

According to study co-author and Institute Fellow, Melina Claussnitzer, Ph.D., "It has become clear that the bulk of disease associated variants are located in the non-coding part of the DNA, where the function of the DNA is largely unknown. Non-coding variants are known to contribute to disease through dysregulation of gene expression. But pinpointing the non-coding variants, which confer this dysregulation remains a major challenge."

The authors applied the analysis to genetic variants associated with type 2 diabetes, one of the most prevalent human diseases. The integration of their computational approach together with several experimental approaches (thereby addressing and proving causality) identified a 2 diabetes variant that promotes disease by interfering with gene regulation and altering fat cell function.

Instead of only considering the conservation of DNA sequences across species, the researchers' computational methodology finds conserved patterns of certain sequences that make up transcription factor binding sites (TFBS) where proteins bind to regulate gene expression. To find these conserved TFBS patterns, the computer uses data about a given region around a gene variant in the human genome, and searches for comparable regions in other vertebrate species. The TFBS pattern conservation of the regions is then scored based on the similarity of TFBS arrangement across species. A high score indicates a high probability that this variant affects the regulation of genes, thereby pointing to the underlying mechanism of a disease.