Comprehensive study using bioinformatics predicts the molecular causes of many genetic diseases

February 9 2010

It is widely known that genetic mutations cause disease. What are largely unknown are the mechanisms by which these mutations wreak havoc at the molecular level, giving rise to clinically observable symptoms in patients. Now a new study using bioinformatics, led by scientists at the Buck Institute for Age Research, reports the ability to predict the molecular cause of many inherited genetic diseases. These predictions involve tens of thousands of genetic disease-causing mutations and have led to the creation of a web-based tool available to academic researchers who study disease. The research is due to be published online in the February 9, 2010 edition of *Human Mutation.*

"We now have a quantitative model of function using bioinformatic methods that can predict things like the stability of the protein and how its stability is disrupted when a mutation occurs," said Buck Institute faculty member Sean Mooney, PhD, who led the research team. "Traditionally people have used a very time consuming process based on evolutionary information about protein structure to predict molecular activity," Mooney said, "I think we're the first group to really quantitatively describe the universe of molecular functions that cause human genetic disease."

The research was done in the contexts of inherited single gene diseases, complex diseases such as cardiovascular and developmental disorders and mutations in cancerous tumors. The study focused on amino acid

substitutions (AAS), which are genetically driven changes in proteins that can give rise to disease, and utilized a series of complex mathematical algorithms to predict activity stemming from the mutations.

As a first step, researchers used available databases of known sites of protein function and built mathematical algorithms to predict new sites of protein function said Mooney. They then applied the algorithms to proteins that have disease-associated mutations assigned to them and looked for statistical co-occurrences of mutations that fell in or near those functional sites. Because the computer algorithms are imperfect, researchers compared that information against a data set of neutral AAS, ones that don't cause human disease, said Mooney. "We looked for statistical differences between the percentage of mutations that fell into the same functional site from both non-disease and disease-associated AAS and looked to see if there was a statistically significant enrichment or depletion of protein activity based on the type of AAS . That data was used to hypothesize the molecular mechanism of genetic disease," said Mooney.

Mooney says 40,000 AAS were analyzed which represents one of the most comprehensive studies of mutations. Describing the results, he used the analogy of a car as a protein -- a big molecular machine. "We are predicting how this machine will break down," said Mooney. "We've known the car isn't working properly because it has some defect; now we can hypothesize that the symptom stems from a broken water pump."

The web tool, designed to enhance the functional profiling of novel AAS, has been made available at http://www..mutdb.org/profile. Mooney identified three different areas of research that could be furthered by use of the tool. Scientists who manage databases of clinically observed mutations for research purposes could develop hypotheses about what those mutations are causing on a molecular level;

they may also be able to use the tool to correlate molecular activity to the clinical severity or subtype of a disease. Mooney says cancer researchers re-sequencing tumors could use the tool to identify mutations that drive the progression of the malignancy. He also expects non-clinical researchers who work with mutations in proteins to use the tool to gain insight into what is causing the mutations. "We are happy to collaborate with scientists, to share data and help them better identify hypotheses about the specific mutations they might be interested in," said Mooney.

Provided by Buck Institute for Age Research