Species diversity helps researchers refine analyses of human gene mutations

September 3 2009, by Richard Harth

In the new era of personalized medicine, physicians hope to provide earlier diagnoses and improve therapy by evaluating patients' genetic blueprints. But, as a new bioinformatics study emphasizes, the first step must be to correctly decipher the deluge of information locked in our DNA and determine its impact on human health.

In the September issue of Genome Research, Dr. Sudhir Kumar led a team of researchers at the Biodesign Institute at Arizona State University in examining DNA mutations from both healthy and diseased patients. Their work evaluates the reliability of computer models aimed at predicting the eventual effect of such mutations.

Along with Kumar, director of the Biodesign Institute's Center for Evolutionary Functional Genomics, others involved with the study were co-authors Michael P. Suleski, Glenn J. Markov, Simon Lawrence, Antonio Marco and Alan J. Filipski.

Kumar's team focused on single DNA mutations—changes to a person's genome that can sometimes make the difference between robust health and debilitating illness. The current study focused on one specific type of DNA mutation—a single change at a given location along the length of DNA—that alters the resulting protein. These protein changes are the source of much of our individuality, coding for differences such as eye and hair color. Scientists have discovered that each person's genome contains thousands of such protein changes. Other single mutations, however, are linked with severe illnesses like cystic fibrosis.

While experimentation on the enormous number of mutations across human populations is impractical due to volume and cost, Mother Nature, as Kumar points out, has already done an experiment for us, presenting scientists with a set of benign mutations for each protein. The branch of science known as comparative genomics takes advantage of the genetic information collected from the diversity of life on Earth.

We now know that humans display a striking degree of genetic similarity with many other species, particularly non-human primates like chimpanzees, gibbons, gorillas and orangutans, with whom we share over 98 percent of our genes. "Comparative genomics provides the first clues as to what a mutation might mean," says Kumar. "This is an area that is going to become center stage in personal genomics and medicine."

Techniques in this rapidly expanding field make use of existing Web databanks such as GenBank, which contains more than 100 billion DNA and protein sequence elements collected from all walks of life. "These databases already contain the outcomes of nature's experiment, which we can harness by using bioinformatics," says Kumar.

DNA medicine typically uses a suite of computer tools to assess whether a newly discovered protein change is potentially disease-causing or benign. Kumar's study tested the reliability of two of the most widely-used tests, known as SIFT and PolyPhen, by examining over 20,000 mutations from both diagnosed patients and healthy individuals. The results demonstrate that these tests make false predictions of risk up to 40 percent of the time, a rate of reliability that renders them impractical for clinical use.

The objective of the study was to identify where SIFT and PolyPhen tend to fail and where their predictions appear to be more reliable. To accomplish this, Kumar's group examined the proteins in 44 species, from frogs to fish, chimps and gorillas. His group discovered that benign

mutations tend to occur in regions of the genome that allow variation over evolutionary time across species. In these regions, it is easier to make accurate predictions of benign mutations.

In contrast, DNA information essential for life is persistent from species to species. Many DNA positions permit no change over evolutionary time in order to preserve proper function—mutations here would likely be damaging. Reinforcement of this theory was found in the subset of mutations discovered in disease-associated genes. Such mutations are clustered in positions of the genome that are conserved over evolutionary time or in mutant protein sequences that are rarely seen. Amazingly, less than 10 percent of known single gene disease mutations are ever found in other species.

As Kumar notes, evolution has provided researchers with a storehouse of genetic mutations, many of which will prove benign for human health. "Suppose you had a mutation at a certain position," he explains, "and your dog has the same change as you have. It's most likely that that change is not harmful." By the same token, if no other species contains the mutation found in one's genome, it calls for further investigation.

Kumar stresses that it will take a combination of additional DNA sequencing data and improved understanding of protein function to refine the power of computer analyses. In the meantime, his bioinformatics evaluations of current computer tools suggest where such tests may be appropriately used for diagnosis with higher confidence and where their results are more likely to be unreliable.

With the costs of rapid DNA sequencing plummeting, individual genetic profiling is already becoming popular, offering every patient access to an enormous treasure trove of medically-relevant information. According to Kumar, the ultimate challenge will be sorting out what all this genetic information implies for each individual's prognosis. Only then will the

promise of personalized medicine be fully realized.

Source: Arizona State University (news : web)