Computational genetics identifies disease-causing mutations

With 3 billion letters in the human genome, it seems hard to believe that adding a DNA base here or removing a DNA base there could have much of an effect on our health.

In fact, such insertions and deletions can dramatically alter biological function, leading to diseases from autism to cancer. Still, it is has been difficult to detect these mutations. Now, a team of scientists at Cold Spring Harbor Laboratory (CSHL) has devised a new way to analyze genome sequences that pinpoints so-called insertion and deletion mutations (known as “indels”) in genomes of people with diseases such as autism, obsessive-compulsive disorder and Tourette syndrome.

DNA insertions and deletions vary in length and sequence. Each indel can range in size from one DNA letter to thousands, and they are often highly repetitive. Their variability has made it challenging to identify indels, despite major advancements in genome sequencing technology.

They are, in effect, regions of the genome that have remained hidden from view as researchers search for the mutations that cause disease. A team of CSHL scientists, including Assistant Professors Mike Schatz, Gholson Lyon, and Ivan Iossifov, and Professor Michael Wigler, has devised a way to mine existing genomic datasets for indel mutations.

The method, which they call Scalpel, begins by grouping together all of the sequences from a given genomic region. Scalpel – a computer formula, or algorithm – then creates a new sequence alignment for that area, much like piecing together parts of a puzzle.

Read the full, original story: A shift in the code: new method reveals hidden genetic landscape