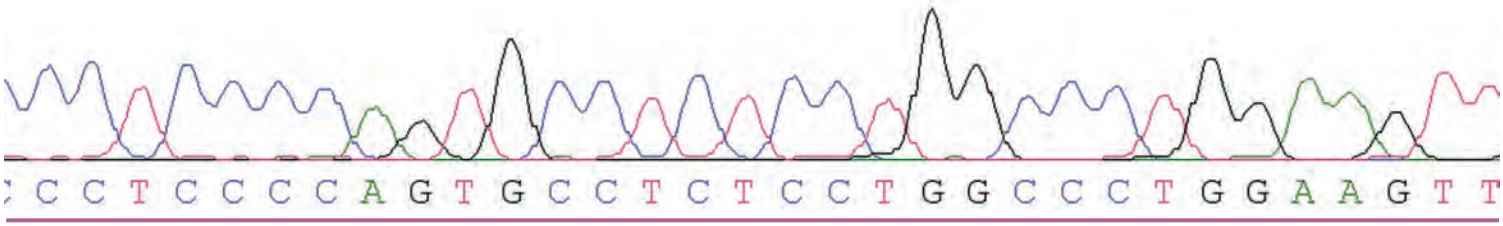
BIOTECH Basics

Genome Sequencing to Uncover Disease



What you need to know:

- *Sequencing the genomes of patients to identify disease-causing mutations is an emerging technology.*
- *A small number of early successes have been recently publicized, providing a glimpse into the possibilities of genome-based medicine.*
- *There are many challenges associated with this type of analysis, including how to correctly interpret the enormous number of genetic changes identified, what information should be disclosed to patients and if this currently represents the best use of limited medical funds.*

If you want to know more:

http://www.jsonline.com/features/health/111224104.html
One In A Billion*, a Pulitzer-prize winning series by the Milwaukee Journal Sentinel, tells the story of Nicholas Volker through words, photos, animation and video.*

http://www.thetech.org/genetics/news.php?id=142
Scientists Improve Patient Treatment through Personal DNA Sequencing: Human Genome Project Finally Helping Individual Patients. *This article is featured on the* Understanding Genetics *website sponsored by Stanford School of Medicine and The Tech Museum in San Jose, Calif. It tells the story of how genetic sequences were used to guide treatment for a movement disorder shared by a twin brother and sister.*

As a toddler, Nicholas Volker failed to gain weight. Doctors at Wisconsin Children's Hospital discovered his intestines were inflamed and ulcerated. When the boy ate, holes would form in his intestine, spilling the contents into his abdomen. While Nicholas' symptoms resembled Crohn's disease, the usual treatments were ineffective. More than 100 surgeries were performed during the next two years, including the removal of his colon. Still, Nicholas' condition worsened.

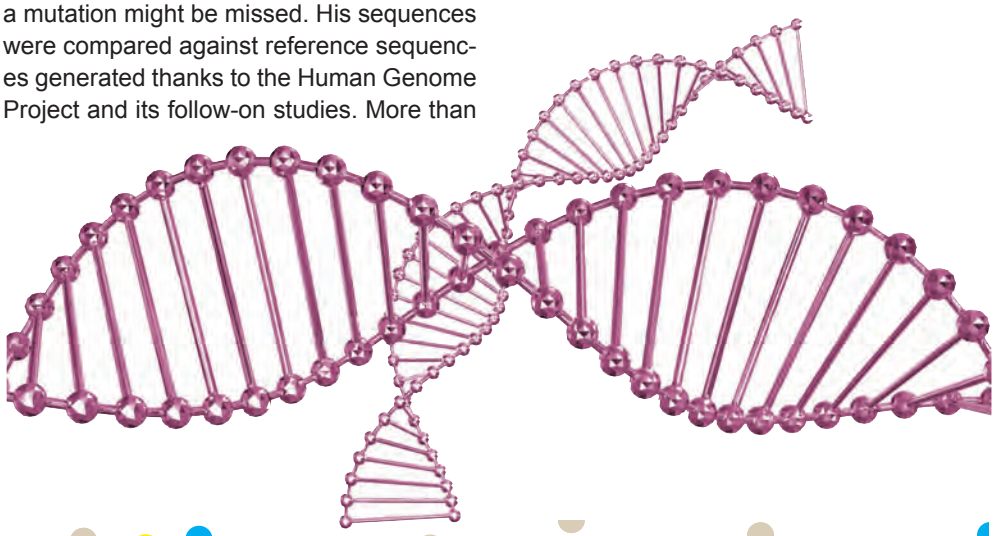
The symptoms suggested a possible immune deficiency and a bone marrow transplant was recommended. Its success, however, depended on identifying the specific underlying cause of Nicholas' symptoms. Consequently, the medical team decided to sequence Nicholas' DNA. They began by sequencing individual genes that were known candidates for irritable bowel syndrome and Crohn's disease. No mutations were identified. The search expanded to sequence the exons of all the genes in Nicholas' genome, a process known as exome sequencing (see table for a comparison of DNA based tests).

Approximately 1 percent of the human genome consists of the genes. Nicholas' exons were sequenced extensively – an average of 34 times, to reduce the chance that a mutation might be missed. His sequences were compared against reference sequences generated thanks to the Human Genome Project and its follow-on studies. More than 16,000 variations in Nicholas' DNA recipe were discovered. Many of these differences had no observable effect. Others influenced physical features like his blond hair and hazel eyes.

After an intensive search through the results, a mutation was discovered in the *XIAP* gene which functions in the inflammation pathway. It was known that mutations in this gene lead to a fatal immune disease, but Nicholas was the first case where the mutation linked to intestinal symptoms. Other researchers working with human genomes were consulted to make sure no healthy individuals had the *XIAP* mutation. In all, more than 2,000 human samples were compared – none showed the genetic change.

XIAP is a gene that sits on the X chromosome. It produces a protein that blocks a pathway leading to cell death. It also helps corral the immune system from attacking the intestine. Nicholas has a single letter change – an A rather than a G. This replaces a cysteine amino acid with a tyrosine at a spot nearly midway though the protein. A set of laboratory tests confirmed that Nicholas' XIAP protein was not functioning as it should.

In addition to the illness, mutations in *XIAP* also cause an extremely rare and potentially lethal disease. The only cure is a



bone marrow transplant. Doctors hoped a transplant would simultaneously treat the illness in his gut. With this information in hand, the 5-year-old underwent an umbilical cord blood transplant from an anonymous donor. The cells in cord blood are similar to those in bone marrow and often are used as a transplant alternative.

The transplant, although not without setbacks, was a success. One year later, Nicholas is an active 6-year-old excited about being back in school, skateboarding and eating vanilla frozen custard. His immune system is still susceptible to infection but he is generally doing very well.

Today, hospitals and research centers across the country are launching their own programs to sequence the exomes, (see table), of select patients. This shiny new tool in a physician's collection comes with a number of challenges and considerations. Initially, a few patients will have their exomes read. Among the first in line are individuals with rare hereditary diseases where it seems likely that a single gene is the root cause. For example, one health care group is sequencing members of a family impacted by congenital hearing loss. The Wisconsin group that sequenced Nicholas' DNA has already sequenced four additional children, with two more in the pipeline.

Once the sequence is analyzed, there is no guarantee that a mutation will be identified or that the results will even influence patient treatment. For complex disorders, it will be difficult to convincingly determine that a specific genetic change contributes to the disease and is not simply part of normal DNA variation.

The sequencing of Nicholas' exons identified more than 16,000 variants. Scientists estimate that sequencing the entire genome may reveal 2 to 5 million genetic changes. Complex software will be needed to identify the needles hidden among the haystack of DNA letters.

Furthermore, as additional genomes or exomes are sequenced, the significance of a variant may change. What is initially thought to cause disease may be benign. Data from this first cohort of patients will need to be frequently reassessed in light of new findings.

Additional questions arise about disclosing genetic information unrelated to the referring symptoms. If a patient undergoes genome sequencing to identify the cause of an autoimmune disorder, should that patient also be told about the genetic mutation that dramatically increases his risk of Alzheimer's? Does the answer differ if the patient is a young child instead of an adult?

This technology is costly. One company that provides whole-genome sequencing services charges $10,000 but reduces the cost to $7,500 when the sequencing is medically justified. This figure does not include the substantial cost involved in analyzing the sequence to identify the relevant DNA changes. That component likely adds tens of thousands of dollars to the final total.

At the end of the day, the story of Nicholas Volker serves as an example of the promise offered by unraveling the genome's secrets. Everyday use of this technology is still down the road, but Nicholas offers an exciting glimpse into the potential impact of genetic information on health and wellbeing.

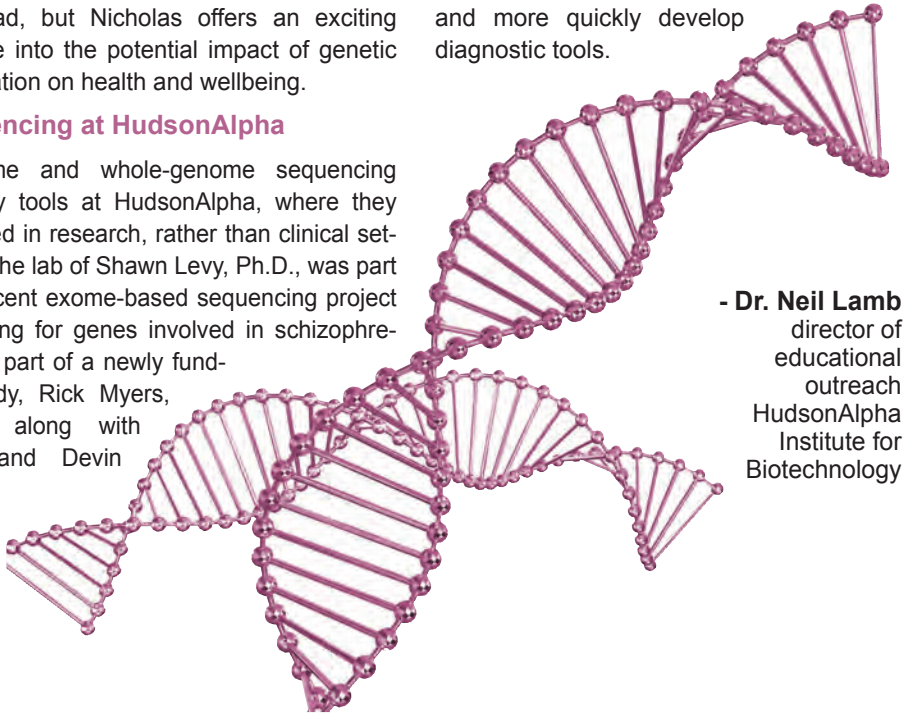
Table: Types of DNA-based tests

Category of Test	Description
DNA identification and profiling	Studies a handful of variable regions of DNA to determine the likelihood that a sample belongs to a particular individual or group. Routinely utilized in forensic studies, this approach does not provide any information regarding disease risk.
Single gene	Generally examines a specific region of a known disease-causing gene, e.g. a frequently mutated DNA nucleotide. Some gene-based tests sequence across the entire gene, including both exons and the intervening spaces known as introns. Several thousands of nucleotides may be analyzed. Examples include HBB testing for the sickle cell mutation and sequencing the BRCA1 and BRCA2 genes for hereditary breast and ovarian cancer. Historically, these tests have involved older sequencing technologies, although a number of next-generation sequencing approaches are in development.
Exome	Selectively determines the sequence of exons, the coding regions of the genome. This represents only a small fraction of the entire genome, meaning less analysis and a lower cost.
Whole genome	Deciphers the complete sequence of an organism. Note: Some small portion of the genome (composed primarily of highly repetitive regions) resists sequencing, so the "whole" descriptor isn't quite accurate.

Sequencing at HudsonAlpha

Exome and whole-genome sequencing are key tools at HudsonAlpha, where they are used in research, rather than clinical settings. The lab of Shawn Levy, Ph.D., was part of a recent exome-based sequencing project searching for genes involved in schizophrenia. As part of a newly funded study, Rick Myers, Ph.D., along with Levy and Devin Absher, Ph.D., will utilize whole-genome sequencing to identify genes involved in the development of bipolar disorder. Methods to analyze whether DNA changes are truly linked to diseases of this sort have been recently published by HudsonAlpha faculty investigator Greg Cooper, Ph.D.

In hopes of better understanding the relationship between our immune system and disease risk, Jian Han, M.D., Ph.D., has announced the creation of the Repertoire 10K Project at HudsonAlpha. The project will use large-scale sequencing to decipher the immune history of individuals with specific diseases to identify disease risk and more quickly develop diagnostic tools.



- Dr. Neil Lamb
director of educational outreach
HudsonAlpha Institute for Biotechnology