

# Getting to the root of genetics

April 17 2012, By Jennifer Chu

---



Manolis Kellis. Credit: M. Scott Brauer

For Manolis Kellis, a deep interest in biology arose partly from an immersion in multiple languages.

Kellis '99, MEng '99, PhD '03, an associate professor of electrical engineering and [computer science](#), spent most of his childhood in picturesque Athens, where he had a view of the Parthenon from his family's balcony. He excelled in school and had a natural affinity for subjects such as math and science.

When he was 12, his family moved to France, and Kellis enrolled in a French-speaking school with his brother and sister — a move that forced the siblings to learn the language, quickly. Kellis found he had to work twice as hard, translating math problems first into Greek to solve them, then translating the solutions back into French. As he grew more

comfortable with the French language, literature and philosophy, he recognized connections between his native Greek and French.

When his family moved again four years later, this time to the United States, Kellis was once again thrown into a completely new language environment. The rapid sequence of [immersion](#) in Greek to French to English taught him to see common roots, compound words and similar expressions among the languages — a skill that became useful in his later study of biology.

“By recognizing the common roots and etymology for words and concepts across cultures, I became interested in evolution,” Kellis says. “The same way that I could trace atoms of inheritance in language, my group is now tracing common genes and functions through evolution.”

## **Biology by chance**

Initially, biology was not part of Kellis’ career plans. As an undergraduate at MIT, Kellis studied computer science, and eagerly took on disparate projects from robotics to image recognition and human motion analysis. Biology came on his radar only by chance.

As a first-year graduate student, Kellis recalls walking into MIT’s student center, where he ran into a friend and fellow student of computer science, who was reading a book on genetics. His friend took him to his lab, where he showed Kellis an algorithm he was developing to assemble the human genome. When Kellis asked to see the data behind the algorithm, the friend opened up a file with pages filled with four letters — A, C, T and G, or adenine, cytosine, guanine and thymine, which are the four nucleotide bases that make up DNA. For Kellis, those pages were life-changing.

“It was a moment of extreme introspection,” Kellis says. “As a computer

scientist, seeing my own genome was like reading the zeros and ones of my hard drive ... I knew that these letters encoded the complete instructions for all cellular processes of the human body, and yet no one knew how to interpret them. Having them all on a computer in front of me though, the path was now open for aligning them, manipulating them, searching for their hidden rules and finding their meaning through computation. So that's what I devoted myself to completely, overnight."

From that moment, Kellis immersed himself in the language of DNA, looking for meaningful patterns in the data. His first project was to compare the genomes of four species of yeast, which at the time was a "daunting" amount of data, he says. Kellis developed computational tools to align characters between the four genomes, identifying patterns of change that allowed him to directly 'read' the set of all genes and regulatory elements. The results, his first paper, were published in the journal Nature.

In 2004, Kellis became a member of the MIT faculty, the Computer Science and Artificial Intelligence Laboratory (CSAIL) and the Broad Institute. Since joining the faculty, Kellis has won numerous awards for his research, including the Presidential Early Career Awards for Scientists and Engineers (PECASE), and Technology Review's Top Innovators under 35. Kellis received tenure in 2011.

## Computing DNA

Today, Kellis heads the Computational [Biology](#) Group at MIT, a group of computer scientists who are using computational techniques to understand the genomic basis of complex human disease. The group collaborates with a vast array of experimental scientists at the Broad Institute, Harvard Medical School and other institutions across the country.

Together, the researchers are sifting through huge genomic datasets, looking for regulatory sequences that determine whether genes are turned on or off. Kellis' team is developing computational methods to efficiently comb through the dense genomic soup of A, C, T and G, to pinpoint key patterns that regulate genes. By identifying such regulatory “motifs,” Kellis hopes to get to the genetic root of diseases such as diabetes, Alzheimer's and cancer — illnesses that may stem from not just one gene or even a number of genes, but a complex interplay of thousands of regulatory elements that control those genes.

His group is also looking at a new “language” of DNA, in the form of chromatin — the combination of DNA and proteins into tight packages that fit inside a cell's nucleus. As DNA is packaged, it loops around nucleosomes, which are made of histone proteins. These histone proteins have long tails that stick out from the packaged DNA, which additional regulators can bind to, altering the use of the genetic code by depositing modifications, or “tags,” that remember the regulatory regions active in each cell type.

“Epigenetic changes write a whole other code on top of DNA, in a different [language](#), that is dynamic,” Kellis says. “That allows each cell to remember its role in the body, and specialize for the functions needed for that role.”

His group is now looking at DNA samples from hundreds of patients and healthy volunteers, to see how regulatory motifs and epigenetic modifications change in their genomes, and how these changes relate to disease.

“The availability of personal genomes, personal epigenomes, disease phenotypes and medical records for thousands of individuals has completely changed medical genomics from a place where getting any dataset was extremely hard, to a place where analyzing the datasets is

often the bottleneck,” Kellis says.

“And for that it helps tremendously to be sitting in an interdisciplinary group of outstanding computer scientists and experimental and medical collaborators who can together really tackle any problem.”

*This story is republished courtesy of MIT News ([web.mit.edu/newsoffice/](http://web.mit.edu/newsoffice/)), a popular site that covers news about MIT research, innovation and teaching.*

Provided by Massachusetts Institute of Technology

Citation: Getting to the root of genetics (2012, April 17) retrieved 2 May 2024 from <https://phys.org/news/2012-04-root-genetics.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--