



## Chunxiao Song

LUDWIG WORKING FROM HOME

“Basically, everything around us can be represented in this very **simple and elegant** way. It seemed like the whole world could be written in simple formulas.”

# The chemical biologist

IN 2008, CHUNXIAO SONG WAS JUST beginning his graduate studies in chemistry at the University of Chicago. It was not going well. A foreign student from China, Song had never traveled outside his country, and the pressure and loneliness were starting to get to him. “Graduate studies, especially for foreigners, can be tough,” Song says.

He was professionally adrift as well. During college at Peking University, Song had majored in organic chemistry. Now, eager to harness chemistry to probe the natural world, he had switched his focus to chemical biology. He was adept at designing reactions to create synthetic molecules without concern for their immediate utility. But in the biological world, chemistry is only useful to the extent that it explains or enables discovery.

A dozen years on, that lost feeling is a pleasantly dim memory for Song, who is now an assistant member of the Oxford Branch of the Ludwig Institute for Cancer Research. In 2019, Song, in collaboration with his Ludwig Oxford colleague Benjamin Schuster-Boeckler, published a study in the journal *Nature Biotechnology* that detailed a greatly improved method for mapping a key chemical—or “epigenetic”—modification made to DNA known as methylation. Epigenetic modifications play a critical role in controlling gene expression, and aberrant methylation across the genome has long been known to be a hallmark of cancer. In 2020, Song and his colleagues launched a company named Base Genomics to commercialize their new technology and apply it to minimally invasive cancer detection.



Song, left, with Ludwig Oxford colleagues Skirmantas Kriaucionis and Benjamin Schuster-Boeckler.

Photo by Monty Rakusen

“Ludwig’s generous funding support, the existing strength of the Oxford Branch in cancer epigenetics and the scientific vision of the Ludwig Institute were all important factors for me to pursue the development of this technology and are essential drivers of high-risk, high-reward projects in my lab to advance cancer diagnostics,” says Song.

### **Inspiration**

Song’s journey to this happy outcome began in 2009, when scientists—including Skirmantas Kriaucionis, who later joined Ludwig Oxford—announced the discovery of a new DNA base, 5-hydroxymethylcytosine, or 5hmC, in human and mouse brains. Up to that point, scientists knew of five main bases, or “letters,” that make up DNA in the genomes of higher organisms. There

are the four canonical ones—adenine (A), thymine (T), guanine (G), and cytosine (C)—plus the product of epigenetic methylation, 5-methylcytosine. “People were calling 5hmC a sixth base,” Song says. “It was very exciting, and people everywhere were racing to understand the biological function of this new base.”

Determined to be one of those people, Song dove headfirst into epigenetics. For Song, the chemical groups involved in epigenetic modifications—found not only on DNA but on its protein packaging as well—were “a wonderland for a chemist to play with in an otherwise bland genome.”

Epigenetic analysis is also crucial to a deeper understanding of cancer. While

genetic information can be used to detect the mutations that drive cancer, it offers few clues about a tumor's tissue of origin or the regulation of its genome. In contrast, epigenetic information is very tissue- and disease-specific, since genes are switched on and off in early fetal development to form individual organs, and epigenetic aberrations are common in cancer.

"If not for that discovery in 2009, I probably would have worked on something else," Song says. "But as soon as I entered the field, I was blown away."

### **The wonders of chemistry**

When Song was 10 years old, he happened upon a school textbook belonging to an older cousin that was filled with seemingly arcane symbols. His cousin explained that the symbols were a kind of shorthand for describing the world. NaCl, for example, was sodium chloride – common table salt. "Basically, everything around us can be represented in this very simple and elegant way," Song recalls. "It seemed like the whole world could be written in simple formulas. That was the first time I saw the wonder of chemistry."

By high school, Song was dragging his mother to the capital city of his province to purchase college chemistry textbooks so he could delve deeper into the subject. He won first-in-class in his province in a national chemistry competition that drew from college-level chemistry.

That gave Song an edge in the national college-entry exam in China. "That first-in-class award gave me an extra 20 points on the national exam. That's a huge, huge boost," Song says—one, in fact, that secured Song a spot in the best chemistry program in China, at Peking University.

In college, Song focused on organic chemistry. "The reactions I worked on were very interesting from a chemistry point of

For Song, the chemical groups involved in epigenetic modifications were "a wonderland for a chemist to play with in an otherwise bland genome."

view, but many wouldn't be useful for a very long time," Song says. "I wanted to change to another area of chemistry where my knowledge could have more immediate use." And so Song, whose second-favorite science was biology, applied to a chemical biology graduate program at the University of Chicago.

### **A detection tool**

Following the discovery of 5hmC in 2009, Song's PhD advisor, Chuan He, tasked Song with devising a way to easily detect the new DNA base. The chemical structure of 5hmC and 5mC are so similar—the two molecules differ by just a single atom—that existing sequencing technologies could not distinguish between the two in the human genome.





Working from home during the pandemic can bring some added benefits.

“There was a chemical modification on the human genome that had never been seen before, but biologists couldn’t sequence it,” Song says. “Chuan He wanted to know, ‘Can we use chemistry tools to detect it?’”

The standard method for detecting 5mC was to design custom antibodies to bind to and flag it on DNA. But that approach fell short with 5hmC modifications, since the modification is far rarer in the genome.

In 2011, Song and his colleagues published a paper in *Nature Biotechnology* detailing a detection method for 5hmC. It involved using enzymatic and chemical reactions to selectively attach a molecular tag to 5hmC modifications, making them easier to spot and target. Their method was eventually

made into commercial kits and is still widely used today.

### **Toward liquid biopsies**

After earning his PhD in 2013, Song moved to northern California as a postdoctoral researcher in the bioengineering lab of Stephen Quake at Stanford University. There, amidst palm trees and perennially mild weather, he continued his efforts to harness epigenetic information for clinical applications, refining “liquid biopsy” tools Quake’s lab was developing that looked for biomarkers in free-floating DNA in the blood.

Liquid biopsies are minimally invasive and altogether less risky for patients. A race is on today to use them as diagnostics that detect very rare bits of DNA shed by

tumors, which encode specific genetic alterations associated with various cancers. “At Stanford, I saw an opportunity to combine what I was doing before, which was epigenetic sequencing, with this cell-free DNA-based liquid biopsy,” Song says. “Before, scientists focused only on changes in the DNA sequence itself—mutations, for example—and ignored all the DNA modifications.”

It wasn't that researchers weren't interested in the epigenetic modifications of cell-free DNA—it was just very difficult to detect them. “Cell-free DNA is present in very minute amounts, and it's highly degraded, so you need a very sensitive method to detect the modifications,” Song explains.

## TAPS

On the strength of his postdoctoral work at Stanford, Song was recruited by Ludwig's Oxford Branch in 2016 as an assistant member. At Ludwig, Song's group has been developing technologies to study how epigenetic modifications to DNA, like 5mC and 5hmC, contribute to cancer. Those technologies could also be applied to develop liquid biopsies for early cancer detection, explore the heterogeneity of tumor cells and elucidate drug resistance mechanisms—all of which are primary goals of Ludwig Oxford.

In their 2019 *Nature Biotechnology* paper, Song, Schuster-Boeckler and their colleagues detailed a novel method for mapping DNA methylation. Called TET-assisted pyridine borane sequencing—TAPS for short—it's less damaging and more efficient than the previous gold standard for mapping 5mC and 5hmC modifications in the genome. Biologists had relied on that method, bisulfite sequencing, for decades. But the approach is extremely destructive, degrading as much as 99% of the DNA in samples. This makes it unsuitable for analyzing cell-free DNA, which is only present in minute amounts in blood.

“Cell-free DNA is present in very minute amounts, and it's highly degraded, so you need a very sensitive method to detect the modifications.”

Bisulfite sequencing can only detect 5mC and 5hmC indirectly, by selectively converting unmodified cytosine to another base, uracil (which is not found in DNA but only used by cells to transcribe genetic information into RNA). This approach is not only inefficient, it also complicates the computational analysis of the data.

TAPS is a two-step process that uses an enzyme called TET to gently convert 5mC and 5hmC to a third modification, 5-carboxylcytosine (5caC), which is then converted to thymine—a DNA base that can be read by ordinary sequencing machines. The Ludwig researchers demonstrated that TAPS can generate more accurate epigenetic sequencing data at a lower cost. They also developed two variations of the technique—TAPS-Beta and CAPS—which can be used to detect 5mC or 5hmC, respectively.

“Humans are diploid, meaning we inherit a genome from dad and a genome from mom. With conventional sequencing technologies, it’s very difficult to distinguish between the two copies because they are so similar.”

#### **One run, two types of data**

Song believes TAPS can replace bisulfate sequencing as the new standard in DNA epigenetic sequencing, and his group is now adapting the technique for various clinical and basic research applications. For example, they are exploring how TAPS might be used to perform single-cell epigenetic sequencing to study biologically significant differences between cell-types within tumors.

In March, Song’s group published a paper in *Genome Biology* describing how TAPS could be combined with other technologies to perform long-read epigenetic sequencing. Until recently, TAPS had only been used to read DNA sequences just a few hundred base pairs long. “In some parts of the genome, where you have repetitive regions and genome

rearrangements, this kind of sequencing does not work very well,” Song says.

But so-called third-generation sequencing technologies are able to read DNA sequences numbering tens of thousands of base pairs in length. “We’ve shown it’s possible to combine TAPS with third-generation technologies so we can do long-read epigenetic sequencing as well,” Song says.

This combination opens up new research possibilities. Not only will it allow researchers to map previously unmappable stretches of the epigenome, it will also enable them to study allele-specific methylations more easily. “Humans are diploid, meaning we inherit a genome from dad and a genome from mom,” Song explains. “With conventional sequencing technologies, it’s very difficult to distinguish between the two copies because they are so similar. But with long-read sequencing, we can actually distinguish between paternal and maternal genomes.”

Song believes scientists have only scratched the surface of what TAPS can do. He and Schuster-Boeckler are now exploring a way to collect both genetic and epigenetic information using TAPS. “If you remove the changes to the genome made by the TAPS chemistry and then use the dataset like you would normal whole-genome sequencing, you could use it for genotyping to identify mutations in cancer,” Schuster-Boeckler says.

The goal is to obtain, simultaneously, from one TAPS run, information about not only where the mutations are, but also about the epigenetic state of a sample.

“People and companies are now realizing that having just the genetic information is no longer enough,” Song says. “You need the epigenetic data as well.” ■



Photo by Monty Rakusen