

Profile of Steven Henikoff

Molecular biologist Steven Henikoff's career was born out of a childhood hobby. He loved photography. Not so much the "shutter" part, he explains, though that was fun, but the chemistry part in the darkroom, where the discovery process occurred. Henikoff's hobby sparked a cascade from his undergraduate major in chemistry to his graduate degree in biochemistry, which led to his career chasing a thread of how to understand epigenetic inheritance—the aspects of a cell that are inherited without changes in DNA.

Along the way, he invented several widely used biotechnology tools such as techniques, designed with the help of his wife Jorja, for deciphering the function of protein sequences by using the power of computers. He also overturned some sacred cows of biology, including that there can be genes within genes, an exception to the idea that genetic information is strictly linear (1). And, recently, he discovered an unexpected structure for a nucleosome: the fundamental "packing unit" that DNA wraps around so it can fit inside the nucleus (2).

Henikoff's work earned him a position at the Basic Sciences Division at the Fred Hutchinson Cancer Research Center in Seattle, recognition as an Investigator in the Howard Hughes Medical Institute, and election into the National Academy of Sciences in 2005. In his Inaugural Article published in October 2007 (3), he pulls together insights from diverse studies into an explanation for how chromosomes' centromeres are maintained.

Encouraged to Thrive

Although none of Henikoff's family members was a scientist, he credits his family for his success. He was born and raised in Chicago in the 1950s, the youngest and only boy among three children. His father manufactured and sold plastic furniture covers. Henikoff's mother has been a lover of games and puzzles. "That helped because science is puzzles," says Henikoff.

In addition, Henikoff says his sisters set an intellectual standard that pulled him in their wake. His sisters were star pupils who went on to succeed in business and in clinical psychology. Because they were good in school, teachers expected a lot of him. "I had someone to look up to," he says.

The high expectations made him work hard, and he set his sights on chemistry. In 1964, he entered the University of Chicago (Chicago, IL). Although he locked into chemistry as a major early, his interests quickly wandered. He discovered



Steven Henikoff working on his favorite hobby.

molecular biology and the concept of using chemistry to study life. In his final year at Chicago, the movie "The Graduate" came out. When Dustin Hoffman's character gets cornered by a family friend who tells him that "plastics" are the future, Henikoff imagined himself as a chemist developing new polymers for the plastic covers sold by his father. Instead, he applied and was accepted to Harvard University's graduate program in biochemistry and molecular biology.

The War Intervenes

Although Henikoff had found the field that eventually would be his career, it would take a back seat to political realities of the late 1960s. In spring 1968, graduate student deferments for the draft had just ended. To avoid being drafted, Henikoff joined the Peace Corps. During training in Louisiana, he met Jorja and they quickly were married without asking permission, which got them "deselected" from the Corps because the superiors thought they needed a period of adjustment before being sent overseas. Henikoff took what he thought was a draft-deferrable job as a hemodialysis technician at the University of Illinois Medical Center, also in Chicago.

"I was waiting for the draft to end or for there to be a lottery, and then go to graduate school," he says. "Unfortunately, my appeals to the draft board ran out before any of that happened, and I got selected."

Still, luck was on his side. He spent a year in a clinical chemistry lab at Fort Jackson in Columbia, SC. Then, as he prepared for deployment to Vietnam,

while at Fort Lewis outside Seattle, his superiors randomly assigned a group of soldiers, with Henikoff among them, to go to Germany instead. Jorja joined him there.

"I was stationed near Mainz, where I mostly drove trucks and ambulances until [I was] released in time to enter grad school in 1971," he says. "Although getting drafted with a war going on was pretty traumatic, I have to admit that Jorja and I quite enjoyed the time we spent in South Carolina and Germany. I learned a lot from my work at Fort Jackson, and the GI Bill money helped out a lot in grad school. So I was quite fortunate, all things considered."

In fact, because he learned German in the Army, he managed to place out of his language requirements once he reached Harvard. There, he joined Matt Meselson's laboratory because they were beginning to look at eukaryotic chromosomes, and Henikoff thought it would prove to be interesting.

Thinking Outside the Box

Henikoff found Meselson to be an incredibly creative thinker who got Henikoff started on what would become his life's work. One of the questions they pondered was the C-value paradox, a curious phenomenon whereby the complexity of an organism does not correlate with the amount of DNA in its genome. For example, newts have ≈ 25 times as much DNA as frogs or humans. Meselson reasoned that some of the excess DNA, composed of small repeating sequences called "satellites," might be involved in the emergence of new species.

"That's the kind of intellectual training I had [in Meselson's laboratory]," says Henikoff. "He was someone who would think outside of the box but in a way that was very logical and made a lot of sense."

The problem was that there was no good way to study satellite sequences. For his dissertation, Henikoff turned to experiments using heat shock in what would become his species of choice: *Drosophila melanogaster*, the common fruit fly. Heat shock was a method researchers were just starting to use for turning specific genes on in *Drosophila*. Because heat shock genes are turned on so quickly and to such a high level, researchers could use the method to more readily look at transcription of DNA into messenger RNA

This is a Profile of a recently elected member of the National Academy of Sciences to accompany the member's Inaugural Article on page 15974 in issue 41 of volume 104.

© 2007 by The National Academy of Sciences of the USA

(mRNA) and translation of mRNA into protein.

Henikoff's labmate, Susan Lindquist, was studying the translation process during heat shock, and this drew Henikoff into examining transcription of heat shock genes. He discovered that heat shock also induced transcription of non-protein-coding RNAs, the functions of which still are not understood.

Answering a Call

After Harvard, Henikoff joined Charles Laird's laboratory at the University of Washington (Seattle, WA), planning to continue his work on heat shock RNAs. Laird sparked Henikoff's interest in another phenomenon most researchers thought could not be studied: position-effect variegation. At issue was the finding, discovered in the 1930s, that if a gene is moved close to the chromosome's centromere, its expression often becomes unstable. Sometimes the gene was active and sometimes it was inactive—the genes did not change but their expression did—classic epigenetics.

"All of our DNA sequences are identical," explains Henikoff, "but we're not just blobs of protoplasm. We have different tissues and those tissues have different programs and those programs get inherited."

Position-effect variegation seemed like the ideal way to get at epigenetics because it showed so many interesting features. But at the time, epigenetics was an obscure topic and there were few techniques with which to study it.

Luckily, before Henikoff could spin his wheels, opportunity called. Laird was on sabbatical for a year, and one day, Henikoff received a telephone call. Ben Hall, across the street in the genetics department, was using a new technique to clone sections of DNA by function. His crew was working in yeast, but they wanted to try it on something more complex. Hall asked Henikoff to isolate yeast-free *Drosophila* DNA.

"That turned out to be hard because *Drosophila* eat yeast," says Henikoff. "It took me months, but I did it."

When he brought the DNA to Hall's laboratory, everyone was busy with other projects, so Hall asked Henikoff whether he would like to do the project himself.

"There was no room, except between another postdoc and a graduate student at a seam between the benches," recalls Henikoff. "I got the seam. But I was able to get some work done and I eventually did clone a fly *Gart* gene by function. I got a paper in *Nature* (4) and a job out of it."

"What I learned from Ben and all this was that you can be working on something that's really cool, like position-effect

variegation, and nobody pays attention, but you do something useful and people will be knocking at your door," says Henikoff.

Indeed, he received a job offer in the Basic Sciences Division of the Fred Hutchinson Cancer Research Center in Seattle. This was a dream job: a vibrant research community that offered an opportunity to pursue high-risk, high-gain research carte blanche. He has been there ever since.

"The Hutch does it really well," says Henikoff. "The idea is that you will make fundamental discoveries if you let the science dictate your next step."

Necessity Is the Mother of Invention

Henikoff continued a tradition of doing useful and exciting research. He has made a name for himself as an inventor, designing several techniques that have proved to be invaluable not only in his laboratory, but to researchers around the world.

For example, while working out the structure of the *Drosophila Gart* gene, he was having trouble finding the 5' end. Genes start at a 5' end and stop at a 3' end, but there are long stretches of introns in between that do not code for the protein. It is difficult to know the length of introns, so finding the ends can be tricky.

"I came up with a method for sequenc-

His sisters set an intellectual standard that pulled him in their wake.

ing more efficiently to help me get through the long first intron and map the 5' end," says Henikoff. The strategy breaks long DNA segments into successively smaller pieces, facilitating DNA sequencing. "That got me hooked on making tools," he says.

He used his new tool to find the 5' end of *Gart*, and that led him to another, serendipitous finding and, eventually, to another innovative research technique.

One day, he was staring at the sequence of the *Gart* first intron. "It looked funny," he recalls. "I knew where the coding stuff was because I mapped it on there. The intron was rich in As and Ts, giving the block of text an angular appearance. Exons—the protein-coding parts that end up in mRNA—are more G and C rich, and blocks of text have a rounder appearance. So I'm looking at [the intron] and

there's all this angular text, but closer to one end, I see some round text."

He started playing around with that odd-looking section of the intron, translating it into a protein, and searching against available protein sequence databases to find regions of similarity to known proteins. He discovered that there was an entire, independent, gene nested inside the *Gart* first intron and encoded on the opposite DNA strand (1).

"It turns out to be not uncommon," says Henikoff. "There are a good percentage of genes that are inside one another. But at the time, in the mid-'80s, it was pretty groundbreaking."

Meanwhile, his success in finding the gene inside the *Gart* intron by computationally translating DNA into protein and searching sequence databases to find protein similarities got him hooked on the idea that this computer-based strategy could be a powerful tool in understanding gene function. Computers and programs were still slow, so he would set up a search in the morning on his home computer and hope that when he got home, he would have a hit. "Very frequently I'd find something," he says. "Now everybody does this."

From those early, crude search methods, Henikoff began to think of ways to find the most conserved regions of related proteins to help reveal their function. That is when his wife, a mathematician by training, got involved. They started by constructing a database of conserved protein motifs, called BLOCKS. The Henikoffs then developed a technique to extract alignment "scores" from the BLOCKS DNA database to determine how related two proteins might be. They called their technique BLOSUM, for BLOCKS of Amino Acid SUBstitution Matrix.

"We came up with scores that worked better than anything being used at the time," says Henikoff. Indeed, in their paper describing the BLOSUM technique (5), they showed that BLOSUM picked up a relationship between *Drosophila* and *C. elegans* DNA transposition proteins that was missed using the standard alignment scores. "After that paper, BLOSUM was adopted by NCBI (National Center for Biotechnology Information) and people immediately started using it," says Henikoff.

Recently, Henikoff's group has introduced methods to help researchers take advantage of burgeoning DNA sequence resources. One, called Targeting Induced Local Lesions IN Genomes (or TILLING), identifies DNA mutations that can be used to determine the function of a gene of interest. The Fred Hutchinson Cancer Research Center and other institutions provide TILLING services to people

interested in model organism genetics and crop improvement.

Another technique, called Sorting Intolerant From Tolerant (SIFT), predicts the impact of a mutation on a protein and has become a standard computational tool in human genetics studies. Both techniques originated as thesis project ideas for Henikoff's graduate students: TILLING from Claire McCallum and SIFT from Pauline Ng.

Epigenetics

Along with developing new techniques, Henikoff has spent his research time on studies that would lead him back to the interesting paradoxes he began to ponder during his graduate and postdoctoral years, including the C-value paradox and position-effect variegation. The thread tying these issues together, says Henikoff, is epigenetic inheritance.

One of the most successful avenues of this work focused on histones, the proteins that bind together with DNA to form nucleosomes. There are four core histones: H2A, H2B, H3, and H4. H3 has a variant called H3.3. One of his recent studies concentrated on differences between how H3 and H3.3 nucleosomes are assembled (2).

"There are only four sequence differences between H3 and H3.3," says Henikoff. "We've known that for 20 years, and it's been assumed that the two would be interchangeable." It turns out they are not. In fact, Henikoff's postdoc Kami Ahmad showed that H3 is assembled into nucleosomes when DNA is replicated, and H3.3 replaces H3 in nucleosomes at genes that are being transcribed.

"We proposed that maybe this is the basis for epigenetic inheritance of gene expression," says Henikoff. In particular, he suggests that the replacement of H3 by H3.3 in areas of active transcription may pass from cell to cell and then also encourage that site to remain active in the sister cells. This might be one way to ensure that specific genes stay on from one cell generation to the next.

Henikoff's Inaugural Article (3), however, pursues the idea of epigenetic inheritance from another avenue. He tackles an area of research that has long stumped researchers: the centromere—the "pinched" area of each chromosome that plays a critical role in making sure chromosomes split properly during cell division.

"Centromeres are weird," explains Henikoff. "They're so extraordinarily con-

served in function—if the centromere fails to work, the cell dies—but so variable in sequence." In fact, centromeres are embedded in those simple repeat sequences of DNA, the "satellites" that intrigued Henikoff at Harvard and that are so variable between closely related species.

A key insight into this variability came from work by Henikoff's postdoc Harmit Malik. The nucleosomes in centromeres have an H3 variant, called CenH3, that is found only in centromeres. Malik discovered that CenH3 is rapidly evolving in a way that indicates the operation of a genetics "arms race." In other words, there appears to be competition for which CenH3 variant gets passed to the next generation.

Malik and Henikoff proposed that this arms race takes place during meiosis, the process in which gametes form for passage to the next generation. In particular, they think it occurs during female meiosis within the egg, which allows only one of the four gametes produced to be transmitted to the next generation. Centromeres from the mother compete against centromeres from the father for transmission to the next generation. Competition between centromeres might result in segregation defects, so that CenH3 and other centromeric proteins would evolve to restore meiotic parity between competing centromeric satellites, in a never-ending conflict. In this way, satellite sequence differences would represent the outcome of a perpetual Darwinian competition that ultimately could lead to speciation.

The ever-changing sequence landscape of the centromere might be responsible for the fact that no particular DNA sequence appears necessary for centromere localization, but rather that centromeres form wherever lots of CenH3 nucleosomes are assembled. Henikoff's newest research finds that CenH3 nucleosomes have a surprising structure that provides insights into how they are epigenetically inherited. That research was based on an idea from Henikoff's postdoc Yamini Dalal that perhaps centromeres are different because their nucleosomes are different from nucleosomes in the rest of the chromosome, which consist of "octamers": an eight-unit bundle made up of two each of the four core histones.

"This was heretical thinking," says Henikoff. "For over 30 years, we've known—it's been part of our psyche, in all the textbooks—that our whole genome, our chromosomes, are filled up

with nucleosomes and they're all octamers. It's really what eukaryotic biology is built upon. What we're saying is, 'not quite.' We think that the centromeric nucleosome is even older than the octameric one and that it's not an octamer. It's equivalent to half of that."

In fact, after years of "classical hardcore" biochemistry and ruling out any other option, Dalal and Henikoff concluded that nucleosomes in *Drosophila* centromeres are "hemisomes" made up of four histones: one molecule each of CenH3, H4, H2A, and H2B.

In Henikoff's Inaugural Article, he, Dalal, and their colleagues lay out the biological implications of this insight. "We show that the hemisome not only makes sense, but can explain a lot of paradoxical observations people have made over the years," says Henikoff.

Henikoff points out that this finding may explain how nucleosomes evolved. Octamers probably came from the joining of two tetramers, made up of each of the four histones, similar to what are found in centromeres, he explains. Add to that the understanding that Archaea—a prokaryote that eukaryotes likely branched off from early in evolution—have nucleosomes in the form of histone tetramers.

"So it makes us think that since Archaea have tetramers and eukaryotes have tetramers in centromeres, that this is the ancestral form from which the octameric nucleosome evolved," says Henikoff.

Happy with His Hobby

Using new tools and epigenetics to solve questions that have intrigued him since graduate school is pretty cool, Henikoff says. But, as always, he is interested in doing something useful. Some of his colleagues at The Fred Hutchinson Cancer Research Center are excited by his findings, and he hopes that some of his basic research can have implications for cancer treatment.

"Cancer is largely epigenetic," explains Henikoff. "It's a change in the regulation of genes. For example, tumor suppressor genes are typically suppressed epigenetically. And if what underlies cancer is epigenetic, then it may be reversible."

He plans to stay at The Hutch to see what happens. He and his wife are happy in Seattle, where they hike and ski and "all those other Northwest things." But when asked what his hobbies are today, he does not hesitate: "This—research—that's my hobby," he says.

Beth Azar, *Freelance Science Writer*

1. Henikoff S, Keene MA, Fichtel K, Fristrom JW (1986) Gene within a gene: Nested *Drosophila* genes encode unrelated proteins on opposite DNA strands. *Cell* 44:33–42.
2. Dalal Y, Wang H, Lindsay S, Henikoff S (2007) Tetrameric structure of centromeric nucleosomes in

- interphase *Drosophila* cells. *PLoS Biol* Aug;5(8):e218.
3. Dalal Y, Furuyama T, Vermaak D, Henikoff S (2007) Structure, dynamics and evolution of centromeric nucleosomes. *Proc Natl Acad Sci USA* 104:15974–15981.
4. Henikoff S, Tatchell K, Hall BD, Nasmyth KA (1981)

- Isolation of a gene from *Drosophila* by complementation in yeast. *Nature* 289:33–37.
5. Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* 89:10915–10919.