YALE JOURNAL OF BIOLOGY AND MEDICINE 91 (2018), pp.503-505.

YALE JOURNAL OF BIOLOGY AND MEDICINE
YJBM
CELEBRATING 90 YEARS 1928–2018

INTERVIEW

The Evolutionary Origin of the Adaptive Immune System

An Interview with David Schatz, PhD

Jingyao Qiu*

Department of Genetics, Yale School of Medicine, New Haven, CT

Dr. David Schatz is a molecular geneticist/immunologist and chair of immunobiology at Yale School of Medicine. He graduated from Yale University in 1980 with B.S. and M.S. degrees in Molecular Biophysics and Biochemistry. He then pursued his M.A. degree in Philosophy and Politics at Oxford University and his Ph.D. degree at the Massachusetts Institute of Technology supervised by Dr. David Baltimore. Dr. Schatz has led efforts that resulted in the discovery of RAG1 and RAG2 and subsequent insights into their evolutionary origins. Dr. Schatz has received numerous prestigious awards including the National Science Foundation Presidential Faculty Fellows Award, the American Association of Immunologists-BD Biosciences Investigator Award, and election to the National Academy of Sciences. He has also been incredibly active as an editor and reviewer and is strongly committed to enhancing predoctoral and postdoctoral training programs at Yale School of Medicine.

Can you briefly talk about what your group is working on right now? We're especially interested in your work on the evolutionary origins of the RAG recombinase.

We work on the mechanism and regulation of the reactions that assemble and diversify the antigen receptor genes. These are immunoglobulin genes, antibody genes in B cells and T cell receptor genes in T cells. In my graduate work, I led the effort to discover and clone RAG1 and RAG2, which are the recombinase genes that assemble the antigen receptor genes during early B cell and T cell development. Ever since the discovery of the genes, I've been interested in their evolutionary origins. We work on the process of V(D)J recombination, which is mediated by RAG. We're interested in how the RAG enzymes work as an endonuclease, biochemically and structurally. We're also interested in how it is regulated: why it cuts where it does and why it sometimes makes mistakes and targets non-antibody genes.

Back to the evolutionary story. RAG1 and RAG2 were strikingly unusual right from the beginning because they're completely different genes and they encode proteins that have no similarity to one another. When we cloned them, we noticed that they were immediately adjacent to one another. The two genes have no introns in their open reading frames and they are only 6KB apart in the mouse genome. It raised the question of why the two genes were right next to each other. They obviously encode proteins that work together but they don't need to be close together in the genome. They also have an interesting convergent orientation, where they transcribe toward one another. It raised the possibility that they have

*To whom all correspondence should be addressed: Jingyao Qiu, Email: jingyao.qiu@yale.edu.

†Abbreviations: RAGs, recombination-activating genes; RSSs, recombination signal sequences.

Keywords: RAG1, RAG2

Copyright © 2018

come from some fungal element or transposable element. About 10 years later, my group and another group at the NIH made an important discovery, which was that RAG1 and RAG2 together had the ability to act as a transposase to cut a piece of DNA at their recognition sequences, which are called recombination signal sequences (RSSs). They can cut the DNA at RSSs and take that excised piece of DNA and insert it into another target DNA, essentially moving the excised fragment from one location to another. That was an activity that happened in the test tube and it happened fairly efficiently. It had major evolutionary implications because it greatly strengthened the link between RAG and transposons/transposases.

For the last 20 years, we've been thinking a lot about where RAG1 and RAG2 came from. And for a chunk of that time, until about 2005, it was extremely mysterious because what was known about other genomes revealed no RAG1 and RAG2 homologs anywhere except that it was found in all of the jawed vertebrates, from shark to human. RAG1 and RAG2 are always right next to each other in convergent orientation. And then it seemed like they just disappeared, and they weren't anywhere else. This left their evolutionary origin very mysterious. Then it has been a series of discoveries in the past 12 years starting from 2005, with a series of transposable elements that have been discovered in other genomes that are outside the jawed vertebrates. That's been very exciting and has now led us to understand where RAG1 and RAG2 came from. Now it's very clear from recent discoveries that they are derived from an ancient transposable element. Now what we're trying to do is to recreate the evolutionary trajectory of how RAG1 and RAG2 came to evolve from a transposable element, and what were the elements and steps that led to the creation of our current adaptive immune system and thus the recombination system.

David Schatz, PhD, Waldemar Von Zedtwitz Professor of Immunobiology and Professor of Molecular Biophysics and Biochemistry, Chair of Immunobiology, Yale School of Medicine, New Haven, CT.

Your lab led the discovery of the ancient transposable element, ProtoRAG, from which RAG1 and RAG2 were believed to derive from. Can you expand a little bit on that very exciting story?

There is actually an important paper that predates the ProtoRAG paper, which was published in 2016 [1]. Earlier that same year, we published a paper in *Genes & Development* where a very talented graduate student in my lab, (Lina) Marcela Carmona was the first author. We analyzed a transposable element named Transib. Transib is a very widely dispersed, ancient transposable element. What is interesting is that it contains a single open reading frame that looks like RAG1. What Marcela was able to provide evidence for was that Transib may well have been the precursor of what we refer to as the RAG transposon. Transib has what are known as terminal inverted repeats at the ends of the transposable element that look like those RSSs that RAG1 and RAG2 work on. But there is no RAG2 or RAG2 homolog in Transib at all. What Marcela provided evidence for was that Transib might well have been able to acquire a RAG2 open reading frame and then the Transib open reading frame and the RAG2 protein could work together. She also provided evidence for them being able to work together in some artificial systems. The earliest step we believe is Transib acquiring a RAG2 open reading frame and giving rise to an intermediate RAG transposon with RAG1 and RAG2.

ProtoRAG was discovered in the genome of a species called Amphioxus, which is an invertebrate Cephalochordate [2]. The ProtoRAG transposon contained every single feature that we had predicted for a transposable element that would be an evolutionary relative of RAG1 and RAG2: it had a RAG1 like gene and a RAG2 like gene, flanked by terminal inverted repeats that looked like RSSs, as well as other provocative features. In the 2016 paper, we demonstrated that there were numerous sequence and functional similarities between the ProtoRAG protein and the RAG protein. They cut DNA in very similar fashion, mechanistically.

This was a very exciting discovery for two reasons. First it strongly supported the model that there was this ancient RAG transposon, because ProtoRAG looks like the clear evolutionary derivative of that transposable element. Second it now gave us a very powerful tool to study an evolutionary derivative of that early ancestor that went

a different route than RAG1 and RAG2 evolutionarily. RAG1 and RAG2, in humans and in other jawed vertebrates, changed from being a transposase into becoming a recombinase. And the big difference is that while both a transposase and a recombinase need to cut the DNA, a recombinase wants to join the ends back together again in a new configuration to give you this gene recombination process. And a recombinase does not want to transpose. Because if it transposes, that can cause genome instability, insertional mutagenesis, and potentially cancer. Whereas a transposase has a completely different biological imperative, which is that it has to cut itself out of one location and has to make sure that it inserts itself into a new location, otherwise it is an evolutionary dead end. Here we have these elements that are related in sequence and have probably diverged from one another 600 to 700 million years ago. We have this wonderful tool to study two enzyme systems that have been operating under these very different evolutionary pressures.

How do you think the discovery of the evolutionary origins of RAG recombinase help us better understand their functions and the adaptive immune system in general?

Some of the central questions for the evolution of the adaptive immune system have to do with how the gene recombination process became established in our genome and it elaborated and diversified. There are now seven different loci that are known to be recombined by RAG1 and RAG2, three of them in the B cell lineage and four of them in the T cell lineage. In terms of the evolution of the adaptive immune system, seeing how RAG1 and RAG2 transform themselves has a number of implications for how the loci were originally generated and the way the adaptive immune system develops now. Seeing how the two proteins have undergone this evolutionary transformation helps us understand how they bind, bend, and cut the DNA and also has a number of implications for their regulation. One of the things that RAG does so well is that it would only cut the DNA when it brings together the two partners that are going to undergo recombination. This is very important for the protection of the integrity of our genome. Seeing how they evolved and we're now making progress in understanding what were the steps that led to RAG1 and RAG2 to become so tightly regulated. It's been very informative in terms of understanding that aspect of the regulation. Then of course, there is the million-dollar question, which is, what led RAG1 and RAG2 to give up on being a transposase. We're making significant headway on understanding what were the adaptations and changes that RAG1 and RAG2 underwent. When you think about how RAG1 and RAG2 bind and interact with DNA, knowing their transposase origin gives us a particularly fertile and useful perspective of their binding and regulatory properties.

Are we aware of any other transposase that has evolved into recombinase besides RAG1 and RAG2?

RAG1 and RAG2 are one of the best studied and the best understood examples of a well-established paradigm, which is transposon molecular domestication. It turns out that there are a large number of transposons and transposases that they encode, that have been domesticated for functions by the host cells. One of the most striking examples is one of the enzymes of the famous Cas system: the Cas1 enzyme is thought to have evolved from a family of transposable elements called casposons. There are a number of transposases where their DNA cutting activity is still integral to what they do. One of the most striking examples is the enzyme in paramecium, oxytricha, and tetrahymena that undergo this remarkable process of genome reduction. They have one large germline genome stored in what's called the micronucleus and they have what's called the macronucleus, where a large fraction of the genome has been cut out and thrown away. This cutting process is mediated by an enzyme system that is derived from a transposable element. And there are enzymes in yeast that control the mating type switch that are derived from transposons. There are now ever-increasing examples as more and more genomes get sequenced. We have now a better and better understanding of just how much of our genome is transposon derived, so the answer is that RAG1 and RAG2 are not alone, but they are probably one of the best understood systems.

REFERENCES

1. Carmona LM, Fugmann SD, Schatz DG. Collaboration of RAG2 with RAG1-like proteins during the evolution of V(D)J recombination. Genes Dev. 2016;30(8):909-17.
2. Huang S, Tao X, Yuan S, Zhang Y, Li P, Beilinson HA, et al. Discovery of an Active RAG Transposon Illuminates the Origins of V(D)J Recombination. Cell. 2016;166(1):102-14.