A Conversation with Andrés Aguilera

INTERVIEWER: LARA SZEWCZAK

Senior Scientific Editor, Cell

Andrés Aguilera is Professor of Genetics at the University of Seville and Director of the Andalusian Centre for Molecular Biology and Regenerative Medicine (CABIMER).

Lara Szewczak: When I think about the work that you've done, RNA control and regulation doesn't exactly come to mind, because it's more focused on genome instability. Where do you think the worlds of genome instability and RNA-based control mechanisms come together?

Dr. Aguilera: Our main interest is to understand the mechanisms of genome instability, but more than 20 years ago we entered into trying to understand a number of proteins and genes whose function was to keep the genome stable. Little by little, we began to realize that this genome maintenance function was very much dependent on transcription. Then we entered into the discovery of R-loops as a main source of the instability that was produced in the cells when these proteins disappeared. This turned out to be an important source of genome instability: The RNA that is being produced during transcription can go back to the DNA from where it came from and form a DNA–RNA hybrid. If the formation of this hybrid is not regulated in a natural way, it would lead to a block of the replication fork on the one hand, or it might create a loop in the DNA that may be subject to attack by different genotoxics. That leads to DNA damage, DNA breaks, which is what leads to genome instability. We realized that the main goal of transcription is to produce the RNA, but it deals with many dangers and cells have to be careful that the RNA that is being produced doesn't have the opportunity to go back and form these DNA–RNA hybrids.

Lara Szewczak: That's the R-loop, right?

Dr. Aguilera: Yes, and this is where the two fields join together. This is why I'm trying to understand this aspect of transcription that, so far, had not been studied that much. Many groups are trying to understand whether there is a special machinery in the cell that tries to control these nascent RNAs to keep them from forming these DNA–RNA hybrids that are going to be harmful for the cell.

Lara Szewczak: If you go back to the cells, what's the evidence that these actually do form? Clearly they can, but what's the evidence that they do and that they contribute to genome instability?

Dr. Aguilera: The DNA–RNA hybrids are present in a number of processes that have natural relevance—for example, in class switch region recombination and in mitochondrial DNA replication. We are talking about unscheduled DNA–RNA hybrids that theoretically shouldn't be formed. It can be tested by different manners, by purifying the DNA from the cell and then doing different treatments with RNase A or RNase H so that you are able to distinguish the kind of structures that it forms. But the most popular way of detecting these hybrids now is via a monoclonal antibody that was developed in the '80s. This antibody is used now worldwide to detect the DNA–RNA hybrids by DNA–RNA immunoprecipitation and by immunofluorescence. This antibody can also recognize double-stranded RNA. Therefore, it is very important that in doing this analysis, we always remove the signal that is recovered by the antibody by using RNase H, because RNase H is known to specifically degrade the RNA moiety of the DNA–RNA hybrid. This is the most popular way of detecting it, but now people are developing other ways. For example, we have fused the DNA–RNA hybrid domain of RNase H to GFP [green fluorescent protein], and this is a way to detect the hybrid directly. Some other people are using mutated RNase H so that when it goes to the DNA–RNA hybrids to eliminate the RNA, it is not able to do that. Then it is used as a marker that the DNA–RNA hybrid exists there.

Lara Szewczak: When you use those kinds of approaches, how frequently do you see R-loops happening?

Dr. Aguilera: That is a very important question: frequency. We really don't know. It's very low. Most analyses are done by genome-wide immunoprecipitation and usually this means that we are dealing with a population of cells. That allows us to see where the hybrids are forming, but if we now try to identify the frequency at a particular site in the genome, it's very low. We have tried to see that in single cells and single colonies. We have gone through 500 different colonies to check in a particular site, and we didn't find evidence for the hybrid being formed. That means that the hybrid is not something occurring very frequently, but it is very important. Why? Because once it is formed it can lead to damage. It's the same as a

double-strand break. A spontaneous break doesn't occur very frequently, but when it happens the consequences are very important. The key question here is that if you remove the machinery in the cell that prevents the formation of the hybrids or that eliminates the hybrids, then the frequency goes up causing high level of breaks; this is going to be harmful for the cell.

Lara Szewczak: What do we know about the machinery?

Dr. Aguilera: I would divide the machinery according to three main functions. The first one covers the proteins that prevent the formation of the R-loops during transcription. These usually are proteins that are involved in RNA processing. Their main functions are probably RNA export and transcription. They have to do with RNA metabolism itself, but collaterally they are preventing the formation of the loop, because if the nascent RNA is packed by proteins into an mRNP [messenger ribonucleoprotein] particle, then that particle is not free to hybridize into DNA. That would be the first body of proteins that prevent the formation.

The second mechanism is: If those hybrids form spontaneously, now you need a machinery to remove them. The best-known one is RNase H, but if you think in terms of transcription you don't want to eliminate the RNA that is just being produced. In human cells, genes can be 600 kb and then spontaneously a short region can form the DNA–RNA hybrid. If this is removed by RNase H, then it will cleave the RNA. You destroy the RNA, and you have to start back again. Therefore, you want to eliminate the hybrid but allow the RNA to stay there.

Third, if hybrids still remain there, the problem comes after cells get into S phase and the replication fork comes and finds the hybrid. There are a number of proteins that are involved in DNA repair: Fanconi anemia, BRCA. When you remove them you see accumulation, but not because you form more hybrids, but because the fork gets stalled there and you are not able to remove them and therefore, the hybrids remain there.

Lara Szewczak: You've outlined these three types of machinery. What are you most excited about right now?

Dr. Aguilera: I'm most excited about the idea that during transcription, we never thought that cells had collaterally developed extra functions—for proteins that are already there—to prevent the formation of R-loops. One of the things that we have found is that these RNA-binding proteins talk directly to chromatin remodelers in a way that, once the RNA is coming out of the RNA polymerase and it binds to these proteins—the ones that we are focusing on, for example, are the THO complex and some others—they talk to the chromatin histone deacetylases as a way to, what we believe might be, transiently close chromatin to prevent the RNA of getting back to the DNA. Because when you have hypoacetylated chromatin you have more compacted chromatin and therefore you avoid hybrids. But at the same time these proteins talk to a known RNA helicase—UAP56 in human cells—that is very abundant and is known to be an RNA

chaperone that works in splicing and other RNA metabolic steps.

Then, we see in these proteins DNA–RNA unwinding ability. When we overexpress these proteins in a number of conditions we are able to eliminate hybrids. We believe that during transcription there is an associated machinery that is working as a kind of checkpoint on the one side, but at the same time as a kind of machinery that prevents the formation of the R-loops. And if these are formed, the helicase will open the DNA–RNA hybrid so that it will allow the transcription to move on. This is a concept that we never thought before, but I think this explains why we started trying to understand the origin of genome instability and we ended trying to understand what happened with the RNA that is coming out during transcription so that the cells have developed this machinery to eliminate the hybrids.

In bacteria, this is probably not a major problem at this level because in bacteria there is cotranscriptional translation; the ribosomes get onto the RNA at the same time this is being produced. Somehow, that also prevents the formation of the hybrid, because this RNA cannot go back to the DNA. In nature, the danger of this nascent RNA going back to the DNA and forming hybrids that would become harmful is there. Evolution has developed this ancillary machinery just to prevent and eliminate them.

Lara Szewczak: I think there have been recent examples in bacteria where there is uncoupling between transcription and translation, so it seems like a place where maybe some new discovery can go back and happen, if you're thinking about this mechanism.

Dr. Aguilera: Absolutely. Actually, in bacteria there are examples of RNA polymerase backtracking, in which you are really dealing with all this and then you get a higher chance of hybrids remaining there. Even then, the mechanics may be different because the RNA polymerase is moving back. There might be different ways of affecting hybrids.

The field is opening and expanding and probably we have to move cautiously now. One of the bigger issues that I'm finding is that there are many RNA helicases that, when you go to in vitro and check activity, most of them open DNA–RNA in vitro but this doesn't mean they act like that in vivo. We have data that suggests that only a few of the proteins are really involved in this. Many others, we have to be cautious. At the moment, we make conclusions trying to transfer what we see in vitro to what we see in vivo, because many of these proteins are RNA chaperones. It's hard to explain that the cells are going to have so many DNA–RNA helicases so that by just removing one you will get the same phenotype as when you remove the other. That means that there is no redundancy, and that is hard to believe. In that case, I believe that all these proteins are working in the making of the mRNP particle and if you don't make them properly you increase the chance of this RNA going back to the DNA.

Lara Szewczak: Roy Parker's talked about potential different ways to think about all the helicases.

Dr. Aguilera: Absolutely. I believe that the major function of those helicases is this one, because RNA helicases have to live with the standard RNAs. But collaterally some of them—UAP56, for example—are so abundant in the cell that, partially, one of the things that they do is to remove the hybrids. You have to think that this protein works cotranscriptionally. You can see them in the DNA. They are recruited to chromatin at the same time there is transcription; they are there during transcription. It makes sense that they might do that job.

Lara Szewczak: You've got the polymerase and all its accessory factors and now you've got all the factors that are binding the RNA and now these helicases that are there sort of on surveillance. It starts to feel like a very crowded environment.

Dr. Aguilera: Absolutely.

Lara Szewczak: You said that you've got an example where something is talking to chromatin modeler. What do you mean by "talking to" and how direct are these interactions likely to be in this very complex soup?

Dr. Aguilera: I believe that it is a complex structure. I don't think it's a direct interaction. We can see that in vivo they happen in the nuclei but when we check other proteins that we know they also interact with—for example, the THO complex—we can see also that they interact. We believe they are in a large complex that is able to interact. We don't know physically who is making the real interaction. We need to explore this farther, but this is—as you mentioned—a crowded structure. Probably, there are different things going on at different times.

For example, one interesting thing we are approaching now is what happens between G₁ and S phase. We are observing that not all proteins are protecting from hybrids in the different stages of the cell cycle. When you remove some of them you see an increase in G₁ and also in S phase, but some others when you remove them you see an increase in hybrids in S phase, but not in G₁. Therefore, some of these proteins may have a specialized function, probably not only in this specific structure but during the

cell cycle. Therefore, we have to deal with many things. When we go into human cells, there is very strong evidence that in neurons hybrids are formed at high levels. Probably what is happening in noncycling cells versus cycling cells is something that we have to explore farther, because the factors that protect one versus another may be different.

Lara Szewczak: When you think about different cell types where you have to worry about R-loops and their biology and the physiology downstream, what are going to be the implications of understanding these mechanisms?

Dr. Aguilera: This is what brings our research to the side of biomedicine. Genome instability is a hallmark of cancer cells. We get many R-loops at the same time leading to genome instability: This is something that we expect to be increased in cancer cells. One thing that we didn't expect was when some years ago we found out that when you remove *BRCA2*, then you increase the R-loops. I think this is a connection that might be somehow present at least in a number of cancers. If that is the case, then certainly the physiology is affected. At the same time, we can consider that this structure might be used as a diagnostic or in some cases even as a target. If the R-loops are enhanced in a number of cancer cells and you get into some molecule that gets into the R-loops and blocks them, you might specifically kill the cells that contain high levels of R-loops versus others. I think there is a good reason to move farther, not only to understand the conceptual part, but the implications.

Lara Szewczak: Your lab focuses a lot on basic research but the idea of a molecule being able to get in either as a diagnostic or as a therapy: Are you dabbling in that space too?

Dr. Aguilera: Absolutely. We are indeed working with companies on a number of tumoral genotoxic agents to see whether they affect better the cells that are accumulating R-loops and whether they target better to regions where the R-loops are accumulating.

Cold Spring Harbor Symposia on Quantitative Biology

A Conversation with Andrés Aguilera

Cold Spring Harb Symp Quant Biol published online December 20, 2019
Access the most recent version at doi:[10.1101/sqb.2019.84.038992](https://doi.org/10.1101/sqb.2019.84.038992)

P<P Published online December 20, 2019 in advance of the print journal.

**Creative
Commons
License**

This article is distributed under the terms of the <http://creativecommons.org/licenses/by-nc/4.0/>, which permits reuse and redistribution, except for commercial purposes, provided that the original author and source are credited.

**Email Alerting
Service**

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To subscribe to *Cold Spring Harbor Symposia on Quantitative Biology* go to:
<http://symposium.cshlp.org/subscriptions>